

5

Resultados das Análises de DIF

Como já vimos, o Funcionamento Diferencial de Itens (DIF) com relação à dificuldade ocorre quando alunos com a mesma proficiência têm probabilidades distintas de acertarem o item. Neste trabalho os diferentes grupos são constituídos pelos alunos brasileiros e pelos alunos portugueses.

No presente capítulo, em um primeiro momento, realizo uma análise exploratória do funcionamento diferencial desses itens do PISA 2003, perseguindo os seguintes objetivos:

- (a) Identificar os itens que se mostraram mais fáceis para o grupo de referência (Brasil) e os itens que se mostraram mais fáceis para o grupo focal (Portugal). Esclareço que no âmbito do presente trabalho, a expressão “itens mais fáceis para alunos de um grupo em relação a outro grupo” sempre pressupõe comparações controladas pela proficiência em Matemática. Em outras palavras, comparo alunos com desempenhos semelhantes no teste.
- (b) Identificar características gerais dos itens para os quais houve DIF e procurar associar essas características com o sentido (pró alunos brasileiros ou pró alunos portugueses) do DIF encontrado.
- (c) Para os itens do PISA que não estão sujeitos a sigilo, discutir o conteúdo dos itens e relacioná-lo com as características gerais dos itens, bem como com o sentido do DIF.

5.1

Investigando os Itens que apresentaram DIF

No início deste estudo, a fim de identificar os itens que apresentaram DIF, utilizei o método da regressão logística, adotando como conhecida a proficiência estimada no PISA. A variável **pv1math** (*Plausible value in math*), presente na base de dados do PISA e calculada para os examinandos dos diferentes países, foi utilizada no modelo de regressão para a estimação do parâmetro ***b*** de dificuldade dos itens.

Para o cálculo das proficiências, o PISA utiliza o modelo da TRI de um parâmetro, criado por Rasch (Rasch,1960)¹ para itens dicotômicos, do tipo certo ou errado, cuja a probabilidade de acertar ($y=1$) ou errar ($y=0$) o item é modelado como:

$$P_i(q) = \frac{e^{D(q-b_i)}}{1 + e^{D(q-b_i)}} \quad (i = 1, 2, 3, \dots, n)$$

Onde:

$P_i(q)$ = probabilidade de um examinando com proficiência q responder corretamente o item i ;

b_i = é o parâmetro de dificuldade do item i ;

n = é o número de itens no teste;

$e = 2,718281828...$

D = constante que vale 1,7.

No modelo de Rasch, o parâmetro b_i (dificuldade) do item corresponde ao ponto, na escala de proficiência q , onde a probabilidade de responder corretamente o item é 0,50. A constante D foi incluída na fórmula para tornar a curva logística igual à curva normal acumulada (ogiva) utilizadas nos estudos pioneiros da TRI.

Note que, pelo modelo de um parâmetro, se um indivíduo tem proficiência q igual à dificuldade do item b_i , a probabilidade de acertar o item é de 0,50 (50%).

$$P_i(q) = \frac{e^{D(q-b_i)}}{1 + e^{D(q-b_i)}} = \frac{e^0}{1 + e^0} = \frac{1}{2} = 0,5$$

No modelo de regressão logística utilizado, a proficiência foi considerada como a única variável explicativa e como variável dependente da resposta dada pelo aluno, que assume apenas dois valores – (1)acertar ou (0)errar - considerando itens dicotômicos. A forma de S que a função logit assume tem a seguinte fórmula:

$$P_i(y) = \frac{1}{1 + e^{-y_i}} \quad \text{onde } Y_i = b_0 + b_1 \text{profic}_i$$

Essa mesma equação pode ser reescrita da seguinte forma:

¹ *One-Parameter (rasch) model* (Rasch,1960) – Technical Report PISA 2003 . pág. 278

$$P_i(y) = \frac{1}{1+e^{-y_i}} = \frac{e^{y_i}}{1+e^{y_i}} = \frac{e^{b_0+b_1 \text{ Profic}}}{1+e^{b_0+b_1 \text{ Profic}}} \quad (\text{a mesma do Modelo de Rasch})$$

Conhecida a proficiência de cada aluno ($pv1math$) e igualando $b_0 + b_1 \text{ Profic}$ a zero podemos calcular o parâmetro de dificuldade do item, que é o ponto da escala onde a probabilidade de acertar o item é de 50%. Assim, $b_i = \frac{-b_0}{b_1}$.

A partir dos b_i encontrados podemos investigar o DIF pela diferença entre os parâmetros de dificuldade do item para os dois países considerados. Essas diferenças nos parâmetros de dificuldade dos 23 itens, calculadas para Brasil e Portugal, constam do apêndice 1.

Depois de calculados os valores acima e verificado que alguns itens apresentavam DIF considerável, foram produzidas as CCI a fim de ratificar o comportamento anômalo de alguns itens. Esses gráficos foram produzidos utilizando o *software* matemático winplot e os valores de b_0 e b_1 estimados na regressão logística.

Por fim, visando obter a magnitude do DIF encontrado, utilizei os *softwares* BILOG-MG 3.0 e SisAni (Sistema de Análise de Itens). O critério aqui adotado foi a estatística de Mantel-Haenszel (MaH), cujos valores e respectiva classificação considerada foram descritos na tabela 2 do capítulo 2.

Essas estatísticas, que servem para identificar se um item apresenta DIF e sua magnitude, foram geradas tendo o Brasil como grupo de referência e Portugal como grupo focal. Os itens que apresentaram DIF e respectivas direções (pró Brasil ou pró Portugal) e magnitudes constam do apêndice 2.

5.1.1

Estimação dos Parâmetros dos Itens e das Proficiências

A análise de DIF, como descrita na seção anterior, foi precedida da estimação dos parâmetros dos itens e das proficiências. Para isso, utilizei o *software* Bilog-MG 3.0. O modelo utilizado foi o logístico de um parâmetro.

Todos os itens da base do PISA-2003 foram recodificados para dicotômicos, tipo certo ou errado. Aqueles que apresentavam escores parciais (*Partial Credit*), para uma resposta meio certa ou uma fase da solução concluída, foram recodificados para errado (*No Credit*). Considerei como resposta correta apenas os que receberam os escores completos (*Full Credit*). O estudo foi feito com toda a amostra do conjunto de dados e precedida de uma análise clássica dos 84 itens, feita no *software* SisAnI. Essa revelou que todos os itens apresentam correlação bisserial alta, acima de 0.45, o que para a Teoria Clássica dos Testes significa bons itens. Para itens dicotômicos, valores elevados nessa estatística significa uma boa discriminação e aumento considerável na confiabilidade do teste. O mesmo resultado, quanto à qualidade dos itens, foi obtido na fase 1 do Bilog. O resultado da estimação dos parâmetros e uma cópia do *output* da fase 2 do Bilog consta do anexo 3. Cabe ressaltar que os parâmetros de discriminação a e acerto casual c , bem como seus respectivos erros-padrão, são constantes em função do modelo de TRI escolhido.

As proficiências estimadas para cada um dos 227.264 alunos constam no *output* da fase três do Bilog e por razões óbvias não serão apresentadas aqui. Não obstante a utilização de um modelo da TRI não considerando os créditos parciais em virtude das recodificações realizadas nos itens, já descritas acima, a correlação entre as proficiências estimadas pelo PISA e as que estimei é bastante elevada: 0,902, em um nível de significância de 0,01.

A análise gráfica dos 84 itens comprova a qualidade destes, utilizados no teste. Em todos os gráficos o principal pressuposto da TRI é respeitado, ou seja, a proporção de alunos que escolhem a resposta correta cresce com o aumento da proficiência e, ao contrário, a proporção de alunos que erram o item decresce com o aumento da proficiência. Os gráficos que descrevem o ajuste do modelo aos dados empíricos, gerados para todos os 84 itens, também apontam para um resultado bastante satisfatório. Ao longo do texto apresentarei apenas os gráficos referentes aos itens públicos. Para os demais itens sigilosos, ver apêndice 3. A seguir apresento um exemplo desses gráficos, relativo ao item M145Q01T.

Gráfico 8: Informações e CCI do Item M145Q01T

Item: M145Q01T

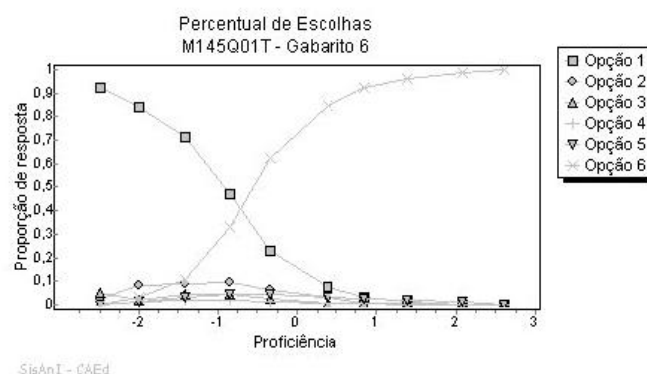
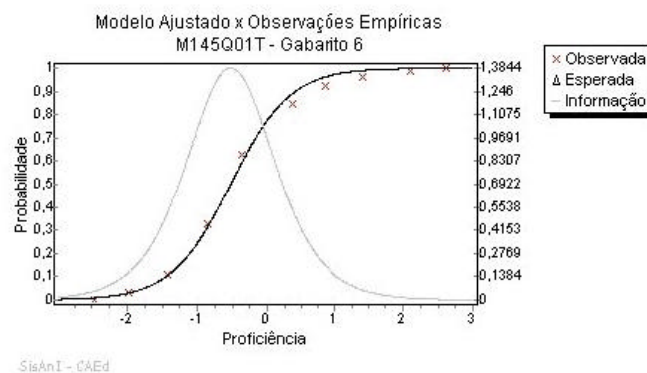
Param. a: 1,38430

Param. b: -0,51699

Param. c: 0,00000

G2: 1028,30000

p-Value: 0,00000



Fonte: SisAni

5.2

Explorando as Características dos Itens Associadas ao DIF

O Quadro 35 identifica os códigos dos itens para os quais houve DIF, indicam as características gerais desses itens e mencionam o sentido do DIF. A operacionalização deste mapeamento foi realizada com base em características definidas pelo próprio PISA, como: “*Subárea da Matemática*”; “*Processo*”; “*Contexto*” e “*Tipo de Resposta*”, e duas outras características, inerentes aos itens, definidas por mim, na expectativa de que possam ajudar na explicação do DIF. São elas: “*Tamanho do Enunciado*” e “*Recurso Gráfico*”, no caso do item vir acompanhado de alguma figura, gráfico ou tabela.

Quadro 35 : Itens que apresentaram DIF entre Brasil e Portugal

Código	Subárea	Competência	Contexto	Recurso gráfico	Tipo	Enunciado	Mais Fácil P/
M179Q01T	Incerteza	Conexões	Público/Social	Sim - gráfico	Aberta	Curto	Brasil
M302Q03	Mudança e Relação	Conexões	Educ./ocupacional	Sim - gráfico	Aberta	Médio	Brasil
M413Q03T	Quantidade	Reflexão	Público/Social	Não	Aberta	Médio	Brasil
M438Q02	Incerteza	Conexões	Público/Social	Sim - gráfico	Fechada	Curto	Brasil
M442Q02	Quantidade	Conexões	Público/Social	Sim - desenho	Fechada	Longo	Brasil
M484Q01T	Quantidade	Conexões	Educ./ocupacional	Sim - desenho	Fechada	Médio	Brasil
M496Q01T	Quantidade	Reflexão	Pessoal	Não	Fechada	Curto	Brasil
M520Q02	Quantidade	Reprodução	Pessoal	Sim - tabela	Fechada	Longo	Brasil
M598Q01	Espaço e Forma	Reflexão	Pessoal	Sim - figura	Fechada	Longo	Brasil
M603Q01T	Quantidade	Reflexão	Pessoal	Não	Fechada	Longo	Brasil
M806Q01T	Quantidade	Reprodução	Educ./ocupacional	Sim - desenho	Fechada	Curto	Brasil
M150Q03T	Mudança e Relação	Conexões	Científico	Sim - gráfico	Aberta	Curto	Portugal
M155Q01	Mudança e Relação	Reprodução	Científico	Sim - gráfico	Fechada	Médio	Portugal
M155Q02T	Mudança e Relação	Reprodução	Científico	Sim - gráfico	Aberta	Médio	Portugal
M273Q01T	Espaço e Forma	Conexões	Educacional/ocupacional	Sim - desenho	Fechada	Médio	Portugal
M302Q01T	Mudança e Relação	Reflexão	Educacional/ocupacional	Sim - gráfico	Fechada	Médio	Portugal
M411Q01	Quantidade	Reflexão	Educacional/ocupacional	Não	Fechada	Médio	Portugal
M413Q01	Quantidade	Reprodução	Público/Social	Não	Fechada	Médio	Portugal
M413Q02	Quantidade	Reprodução	Público/Social	Não	Fechada	Médio	Portugal
M421Q01	Incerteza	Reflexão	Educacional/ocupacional	Não	Aberta	Curto	Portugal
M438Q01	Incerteza	Reprodução	Público/Social	Sim - gráfico	Fechada	Curto	Portugal
M505Q01	Incerteza	Reflexão	Científico	Sim - tabela	Aberta	Médio	Portugal
M547Q01T	Espaço e Forma	Reprodução	Educacional/ocupacional	Sim - desenho	Fechada	Curto	Portugal

Os primeiros itens descritos no quadro acima são aqueles que apresentaram uma facilidade maior para alunos brasileiros, comparados com seus colegas portugueses de mesma habilidade cognitiva (ver última coluna). Onze itens foram mais favoráveis para o Brasil e 12 itens foram mais favoráveis para Portugal. Assim, tomando como exemplo o primeiro item do Quadro 35 (M179Q01T), um dos 84 utilizados na avaliação do PISA em 2003, na área de Matemática, que apresentou funcionamento diferencial, podemos afirmar que:

- a) A competência requerida por este item é da subárea “Incerteza”, ou seja, seu conteúdo está relacionado com os fenômenos e as relações probabilísticos e estatísticos (ver o item na pág. 163);
- b) A dimensão do processo demandado pelo item é “Conexões”, o que exige do aluno a ligação de idéias para resolvê-lo;
- c) A situação na qual os alunos encontram esse problema matemático ou onde são aplicados os conhecimentos relevantes é de contexto “Público/Social”;
- d) O item apresenta, em seu enunciado, um tipo de “Recurso Gráfico” (gráfico);
- e) O item exige do aluno a elaboração de uma resposta ou argumentação;
- f) É um item cujo enunciado é curto;
- g) E, finalmente, este item foi mais fácil para alunos brasileiros, comparados com alunos portugueses que obtiveram o mesmo desempenho no teste.

Cabe ressaltar que, normalmente os itens que apresentam DIF elevados e sistemáticos são identificados e excluídos em pré-testes e análises estatísticas preliminares, antes de utilizá-los para a produção da proficiência do aluno. Portanto, não se espera, em princípio, que encontremos itens com padrões bem definidos associados a existência do DIF. No entanto, ainda assim, alguns itens que exibem algum grau de comportamento diferencial, como os mostrados aqui, podem trazer alguma informação adicional que seja relevante para entender algumas das possíveis diferenças educacionais, entre outras, existentes entre os países analisados.

Na sequência do presente capítulo passo a descrever cada uma das características gerais apresentadas nos quadros acima, buscando associá-las com o sentido do DIF.

5.2.1 – Subáreas da Matemática: as evidências sobre o DIF associadas a “Quantidade” versus “Mudança e Relação”

A tabela abaixo mostra a distribuição dos 23 itens de Matemática que apresentaram algum funcionamento diferencial, tendo o Brasil como grupo de referência e Portugal como grupo focal. A classificação dos itens, para fins de explicação do DIF, é feita, neste momento, em função da subárea do conteúdo de Matemática que o item avalia. Observemos que as diferenças entre a quantidade de itens para cada grupo, dentro das subáreas, são maiores em duas delas: “Quantidade” e “Mudança e Relações”.

Tabela 11: Itens com DIF, segundo a Subárea da Matemática

Característica do Item (subárea)	Nº de itens com DIF Mais fácil para	
	Brasil	Portugal
Quantidade	7	3
Mudança e Relação	1	4
Espaço e Forma	1	2
Incerteza	2	3
Total	11	12

De fato, dez itens são da subárea “Quantidade”. Desses, sete foram mais fáceis para alunos brasileiros. Com relação à “Mudança e Relações”, dos cinco itens que apresentaram DIF, quatro foram mais fáceis para alunos portugueses. Esses dados indicam, num primeiro momento, que alunos brasileiros tendem a ter melhor desempenho do que alunos portugueses quando o conteúdo avaliado refere-se à “Quantidade”, quando comparado com a subárea “Mudança e Relações”. Tal fato aponta um possível caminho na tentativa de encontrar um padrão para explicar o DIF entre esses dois grupos. O êxito nesse empreendimento passa, necessariamente, por uma análise mais detalhada nas características desses itens.

Como já mencionado, nem todos os itens do PISA 2003 são liberados para divulgação. Tratam-se de itens sigilosos, que foram analisados sob o compromisso assumido de que não seriam divulgados. Tal procedimento se justifica em virtude de muitos itens serem reaplicados em avaliações futuras. Os itens públicos, liberados pelo consórcio internacional que administra o PISA, e disponibilizados

no site do Inep², serão discutidos e apresentados exatamente da mesma forma como apareceram nos testes que foram submetidos aos alunos. Para os demais itens, somente as estatísticas e os gráficos serão apresentadas nos apêndices 2 e 4, respectivamente. É importante observar também que, no PISA, muitos itens, a partir de um mesmo enunciado e auxílio gráfico (figura, tabelas, gráficos, etc), se desdobram em mais de uma questão ou tarefa. Por exemplo, a situação-problema M520, apresentada a seguir, se desdobra em três itens: M520q01; M520q02 e M520q03.

Dos sete itens de “Quantidade” que apresentaram DIF favorável aos estudantes brasileiros, quatro são públicos e serão apresentados abaixo. São eles M413Q03; M484Q01T; M520Q02 e M806Q01T. Além de apresentá-los e discutí-los, explicito resumidamente o resultado da análise que identificou o funcionamento diferencial do item para alunos dos diferentes países.

Observemos o item M484Q1T:

ESTANTES

Questão 1: ESTANTES

Para construir uma estante completa, um marceneiro precisa do seguinte material:

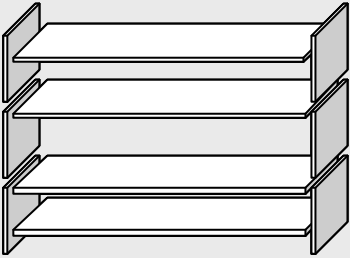
- 4 pranchas grandes de madeira,
- 6 pranchas pequenas de madeira,
- 12 braçadeiras pequenas,
- 2 braçadeiras grandes e
- 14 parafusos.

O marceneiro possui em estoque 26 pranchas grandes de madeiras, 33 pranchas pequenas de madeira, 200 braçadeiras pequenas, 20 braçadeiras grandes e 510 parafusos.

Quantas estantes completas o marceneiro poderá fazer?

Resposta:

M484Q01



A análise gráfica, que normalmente é feita como primeiro passo para a verificação da qualidade dos itens, possibilitando verificar e avaliar, por exemplo,

² www.inep.gov.br

o grau de dificuldade e o poder de discriminação do item, mostra-se um instrumento bastante eficiente, também, para os estudos de DIF. A Curva Característica do Item (CCI), já definida anteriormente, foi utilizada para investigar se esse item apresenta funcionamento diferencial. Como a relação existente entre a probabilidade de acerto no item e a habilidade do respondente é representada graficamente por essa curva, para que um item não apresente DIF é necessário, pois, que a sua CCI seja a mesma para os dois grupos. Caso contrário, podemos afirmar que o item apresenta algum tipo de funcionamento diferencial.

Abaixo apresento a CCI desse item para Brasil e Portugal. Além disso, um “zoom” da região do gráfico de maior interesse também é mostrado.

Gráfico 9: Curva Característica do Item M484q01t

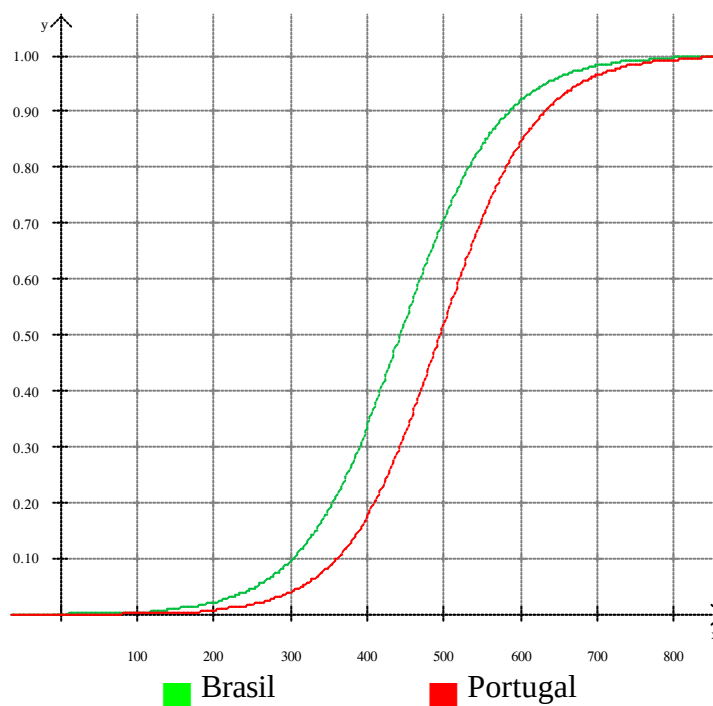
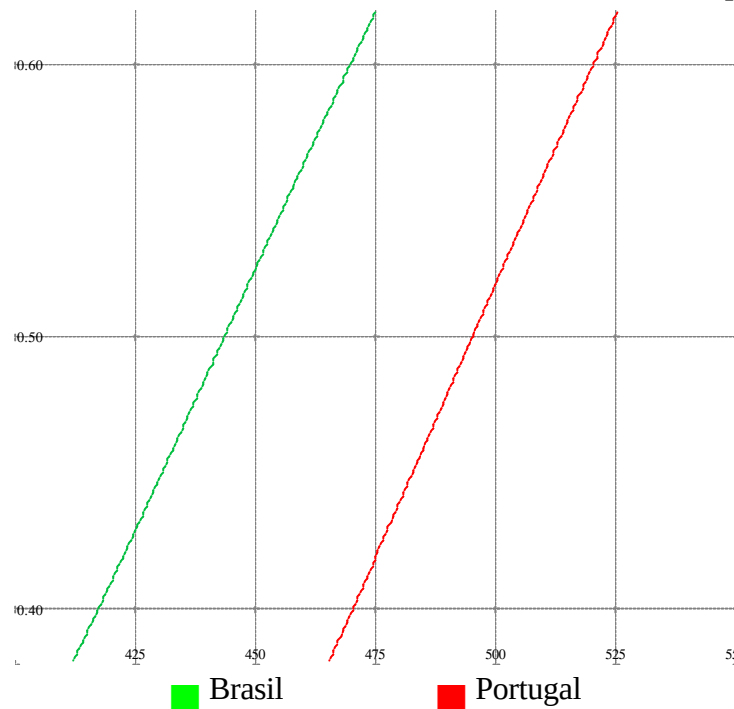
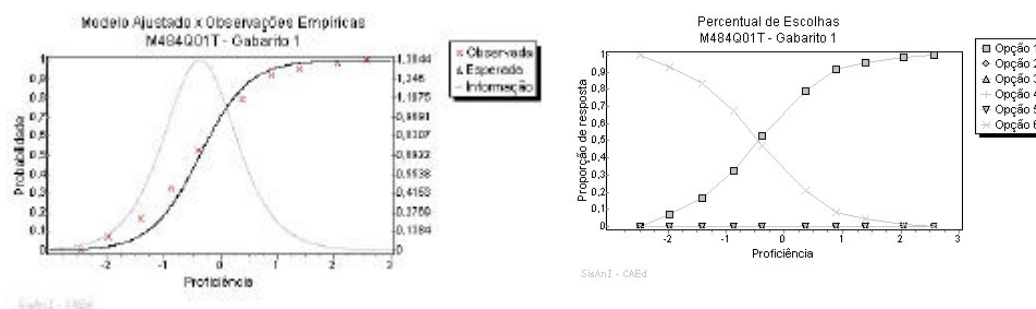


Gráfico 9a: - Zoom da Curva Característica do Item M484q01t



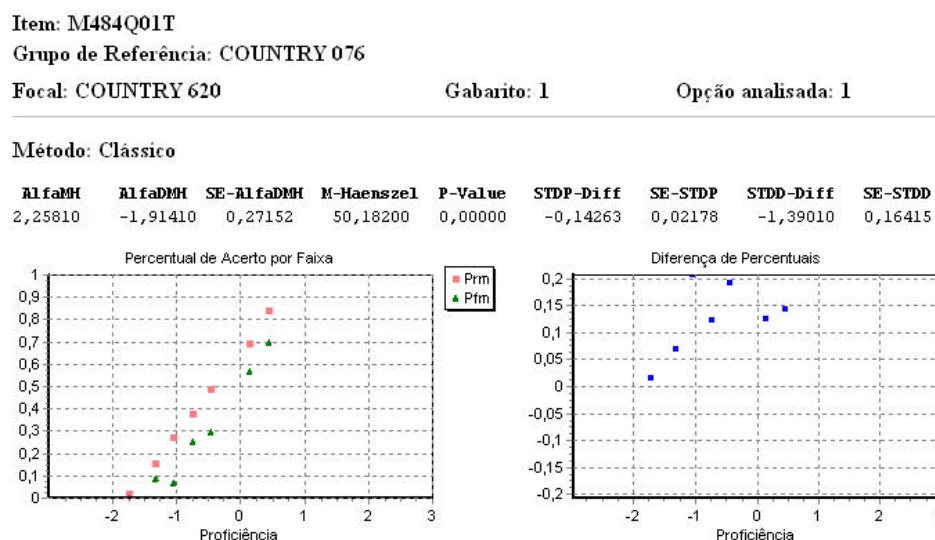
A análise gráfica nos mostra que esse item apresenta DIF apenas na dificuldade (diferença no parâmetro ***b***), pois as curvas têm a mesma inclinação, indicando a não existência de DIF na discriminação (diferença no parâmetro ***a***). Esse é um caso de DIF uniforme, pois o item favorece uniformemente os estudantes brasileiros. Para todas as faixas de proficiência, o item é mais fácil para o Brasil. Observando-se o valor da estatística delta de Mantel Haenszel calculado para esse item ($\alpha D MH = -1,914$), utilizada para analisar a magnitude do DIF encontrado, podemos classificar esse funcionamento diferenciado como de magnitude alta, conforme tabela de classificação apresentada anteriormente. O sinal negativo indica que o item favorece o grupo de referência (Brasil).

Os gráficos abaixo, produzidos pelo *softwares* BILOG-MG e SisAni, dão uma idéia do ajuste do modelo aos dados empíricos e como se comporta o item em relação à escolha pela opção correta e os distratores. No gráfico 10, o “retângulo” da esquerda mostra a CCI do item e as observações empíricas, o “retângulo” da direita mostra que a linha que representa a opção pela resposta correta cresce com o aumento da proficiência e as demais decrescem, exatamente como é esperado.

Gráfico 10: Ajuste do Modelo e CCI - Item M484Q01T

Já o gráfico 11, abaixo, mostra o comportamento do item ao longo da escala de proficiências para os dois grupos. Nele observa-se que, para todas as faixas de proficiências, o item apresenta uma facilidade maior para o grupo de referência – Brasil, representado pelos “quadrinhos” de cor rosa. Este resultado é coerente com a diferença nos parâmetros de dificuldade apresentada no apêndice 1 e com a análise feita via curva característica do item –CCI, acima.

Esclareço que para não sobrecarregar o texto, esses gráficos, para os demais itens, constarão apenas no apêndice 5.

Gráfico 11 – Percentual por Faixa Acerto – Opção Correta

A explicação para o DIF apresentado por este item, favorecendo alunos brasileiros, reside, aparentemente, na demanda pela realização de cálculos simples. Este item pode ser resolvido com a divisão de números inteiros e noções básicas de agrupamento numérico. No entanto, a análise desse e outros itens

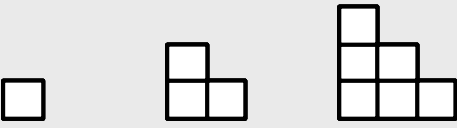
envolvendo a subárea “Quantidade”, exigindo operações aritméticas, necessitam ser vistos em conjunto, pois do contrário haverá perda de poder explicativo para o estudo do domínio de conceitos relacionados entre si, tais como multiplicação, divisão, razão e proporção, pelo fato de que matematicamente eles não são independentes. Portanto, abordaremos essas competências de um modo mais amplo, na perspectiva de Vergnaud (1983), que situa esse tema no campo conceitual das estruturas multiplicativas³. Essa ideia ficará mais clara por ocasião das análises dos demais itens dessa subárea.

O item a seguir é o M806Q01T:

SEQÜÊNCIA EM ESCADA

Questão 1: SEQÜÊNCIA EM ESCADA M806Q01

Roberto constrói uma seqüência com o formato de uma escada, utilizando quadrados. Ele segue as seguintes etapas:



Etapa 1 Etapa 2 Etapa 3

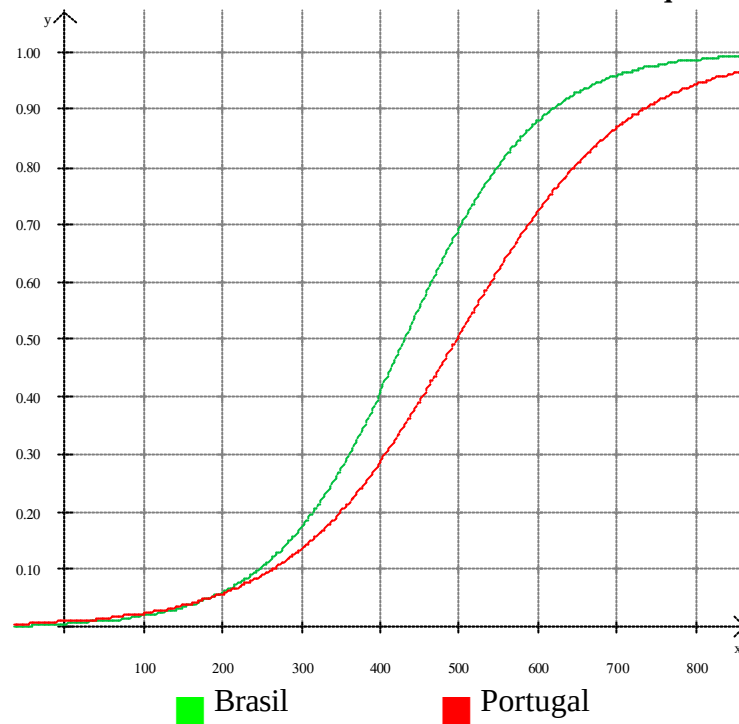
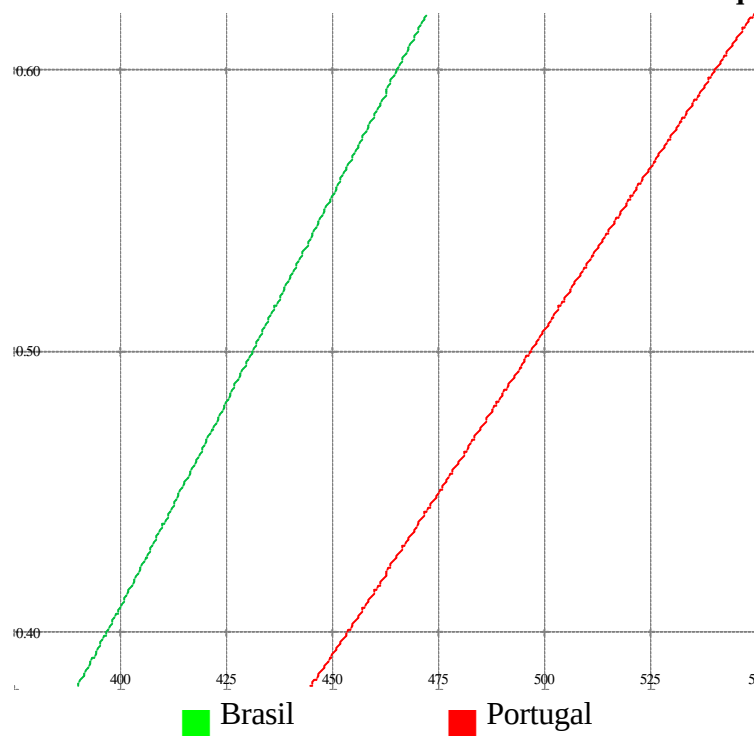
Como você pode ver, ele utiliza um quadrado na Etapa 1, três quadrados na Etapa 2 e seis na Etapa 3.

Quantos quadrados ele vai utilizar na quarta etapa?

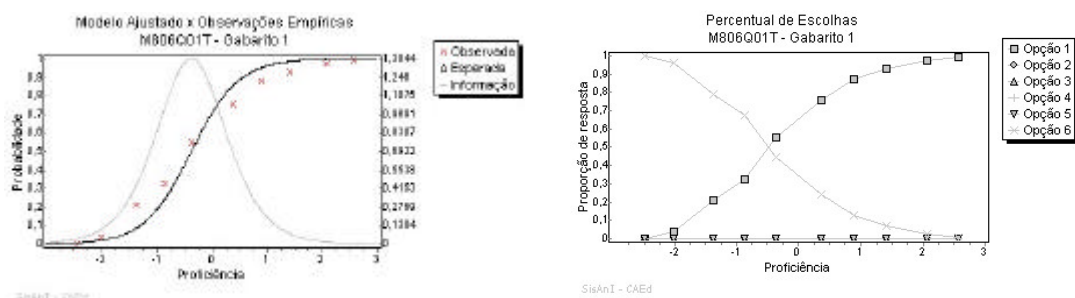
Resposta:.....quadrados.

A CCI desse item, igualmente ao anterior, demonstra a existência de DIF na dificuldade, sendo mais fácil para alunos brasileiros. A estatística delta de Mantel Haenszel calculado para esse item (alfa D MH = -1,437), indica um funcionamento diferenciado de magnitude intermediária.

³ Para Vergnaud, o campo conceitual é “um conjunto de problemas e situações para o tratamento dos quais conceitos, procedimentos e representações de tipos diferentes mas estreitamente interligados são necessários” (1983, p.127). O campo conceitual das estruturas multiplicativas é o conjunto de problemas que envolvem operações e noções aritméticas do tipo multiplicativo, tais como multiplicação, divisão, fração, razão e semelhança.

Gráfico 12: Curva Característica do Item M806q01t**Gráfico 12a: Zoom da Curva Característica do Item M806q01t**

Os gráficos a seguir, mostram um razoável ajuste do modelo aos dados empíricos e um bom comportamento do item em relação à escolha pela opção correta, à medida que aumenta a proficiência.

Gráfico 13: Ajuste do Modelo e CCI -Item M806Q01T

Esse item, da mesma forma que o anterior, favorecendo alunos brasileiros, requer a realização de cálculos simples, podendo ser resolvido com contagem numérica simples.

Na sequência, o item M520Q02:

SKATE

Eric é um grande fã de skates. Ele vai a uma loja chamada SKATERS para verificar alguns preços.

Nesta loja, é possível comprar um skate completo, ou pode-se comprar uma prancha, um jogo de 4 rodinhas, um jogo de 2 eixos e um jogo de acessórios, e montar seu próprio skate.

Os preços dos produtos oferecidos por esta loja são:

Produto	Preço em zeds	
Skate completo	82 ou 84	
Prancha	40, 60 ou 65	
Um jogo de 4 rodinhas	14 ou 36	
Um jogo de 2 eixos	16	
Um jogo de acessórios (rolamentos, amortecedores de borracha, parafusos e porcas)	10 ou 20	

Questão 2: SKATE

M520Q02

A loja oferece três tipos de pranchas diferentes, dois jogos de rodinhas diferentes e dois jogos de acessórios diferentes. Há apenas uma opção para o jogo de eixos.

Quantos skates diferentes Eric poderá construir?

- A 6
- B 8
- C 10
- D 12

A análise gráfica nos mostra que esse item apresenta DIF também na discriminação. As curvas têm inclinações diferentes, mostrando que esse item discrimina mais os alunos portugueses do que brasileiros (a curva de Portugal é mais vertical). Esse é um caso de DIF não-uniforme, pois há uma interação entre o nível de proficiência e a performance no item, ou seja, a direção do DIF muda ao longo da escala. Para alunos com proficiência abaixo de 520, aproximadamente, o item favorece os estudantes brasileiros. Para alunos com proficiências maiores, o item passa a ser mais fácil para alunos portugueses. O valor da estatística delta de Mantel Haenszel calculado para esse item ($\alpha D_{MH} = -1,045$), indica um funcionamento diferenciado de magnitude intermediária.

Gráfico 14: Curva Característica do Item M520q02

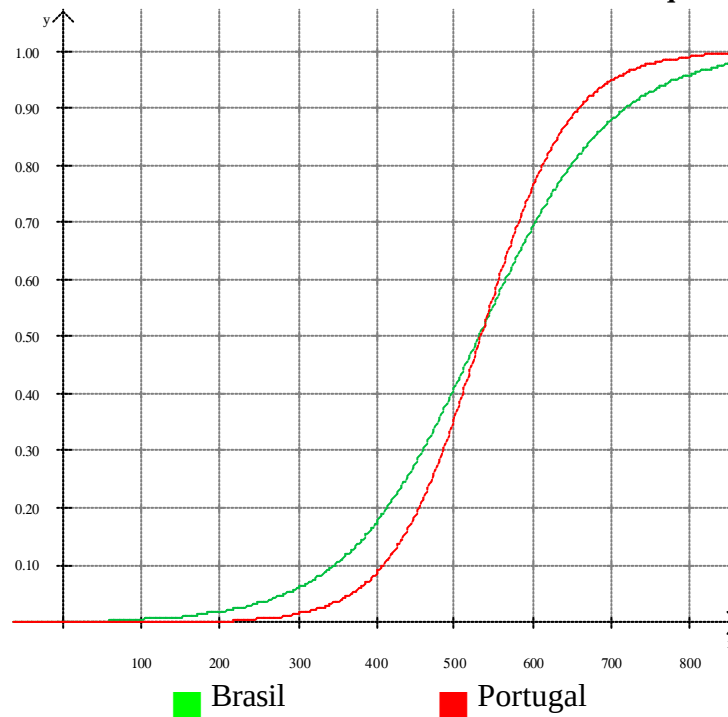
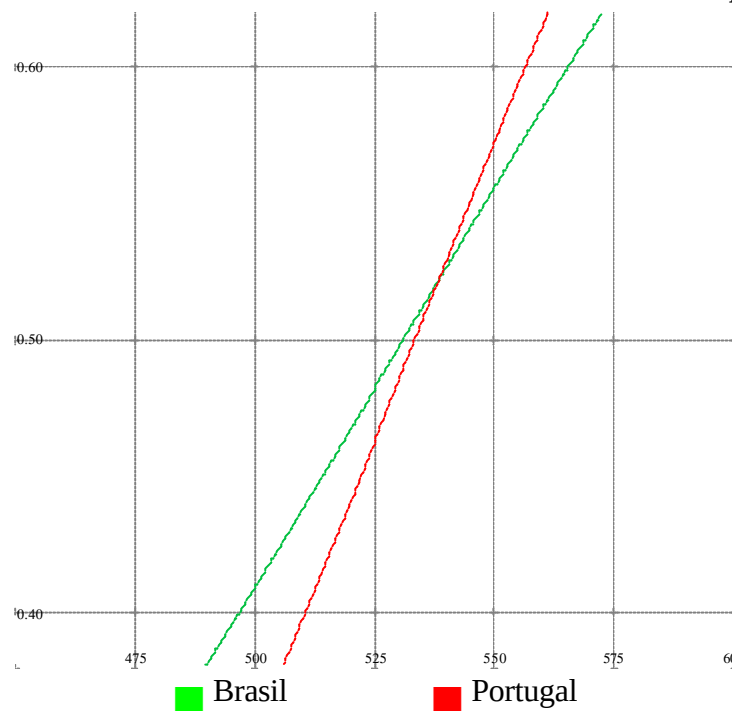
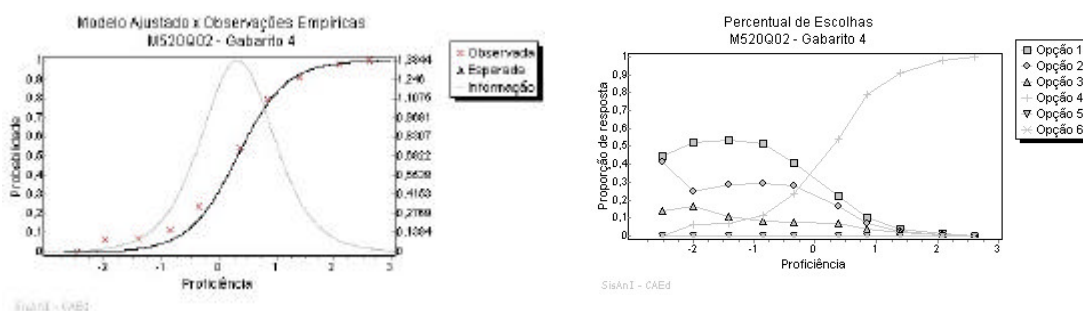


Gráfico 14a: Zoom da Curva Característica do Item M520q02

Como o item é de Múltipla escolha, é possível que o DIF favorável ao Brasil, na região de proficiências mais baixas, se deva ao fato de que os alunos brasileiros se aventurem mais na tentativa de acerto casual (“chute”) do que alunos europeus. De qualquer forma, esse item trata de um tema que, há pelo menos dez anos, vem sendo bastante recomendado pelos avaliadores do PNLD (Programa Nacional do Livro Didático) e já está presente em muitos livros adotados. Também os resultados do SAEB demonstram que essa demanda é bem sucedida por estudantes brasileiros. A competência envolvida para a resolução correta desse item é o princípio multiplicativo.

Os gráficos mostrando o ajuste do modelo aos dados empíricos e o comportamento do item em relação à escolha pela opção correta, são apresentados abaixo.

Gráfico 15: Ajuste do Modelo e CCI -Item M520Q02

O gráfico da esquerda mostra que o desajuste do modelo é mais acentuado na região da curva de menores proficiências e o gráfico da direita mostra que até certo nível de proficiência os distratores atraíram a opção de muitos alunos.

O próximo item a ser considerado é de importância fundamental para a hipótese levantada a cerca da subárea “Quantidade”, na explicação do DIF favorável ao Brasil. Trata-se da situação-problema M413, que se desdobra em três itens: M413Q03T, com DIF favorável ao Brasil e os itens M413Q1 e M413Q2, com DIF favorável a Portugal. Outro item, o M411Q01, não público, da mesma subárea, também apresentou DIF favorecendo os alunos portugueses.

Retornando a idéia de campo conceitual defendida por Vergnaud (1983), todos esses itens da subárea “Quantidade” envolvem operações e noções de Aritmética. No entanto, embora o conjunto de problemas e situações exija conceitos que estão interligados, os tipos de procedimentos e representações são diferentes. De fato, em todos os itens favoráveis ao Brasil, a representação numérica é a mesma: apenas números naturais. Já os três itens de “Quantidade” que apresentaram DIF na direção oposta, ou seja, foram favoráveis aos alunos portugueses (M413Q01, M413Q02 e M411Q01), envolvem as mesmas operações aritméticas, mas avaliam competências relacionadas com números decimais.

O M413Q03, favorável ao Brasil, não necessita da realização de cálculos (com números decimais) para uma resposta correta.

Vejamos os itens M413Q01, M413Q02 e M413Q03:

TAXA DE CÂMBIO

Mei-Ling, de Singapura, estava preparando-se para uma viagem de 3 meses à África do Sul como aluna de intercâmbio. Ela precisava trocar alguns dólares de Singapura (SGD) por *rands* sul-africanos (ZAR).

Questão 1: TAXA DE CÂMBIO

M413Q01 – 019

Mei-Ling descobriu que a taxa de câmbio entre o dólar de Singapura e o *rand* sul-africano era:

1 SGD = 4,2 ZAR

Mei-Ling trocou 3000 dólares de Singapura por *rands* sul-africanos a esta taxa de câmbio.

Quantos *rands* sul-africanos Mei-Ling recebeu?

Resposta:

Questão 2: TAXA DE CÂMBIO

M413Q02 – 019

Ao retornar a Singapura após 3 meses, Mei-Ling ainda tinha 3 900 ZAR. Ela trocou novamente por dólares de Singapura, observando que a taxa de câmbio tinha mudado para:

1 SGD = 4,0 ZAR

Quantos dólares de Singapura Mei-Ling recebeu?

Resposta:.....

Questão 3: TAXA DE CÂMBIO

M413Q03 - 01 02 11 99

Durante estes 3 meses, a taxa de câmbio mudou de 4,2 para 4,0 ZAR por SGD.

Foi vantajoso para Mei-Ling que a taxa de câmbio atual fosse de 4,0 ZAR em vez de 4,2 ZAR, quando ela trocou seus *rands* sul-africanos por dólares de Singapura? Dê uma explicação que justifique a sua resposta.

As CCIs desses três itens indicam a presença de DIF uniforme, sendo os dois primeiros favoráveis a Portugal. As estatísticas delta de Mantel Haenszel, são respectivamente: alfa D MH = 1,454, alfa D MH = 1,421 e alfa D MH = - 2,397.

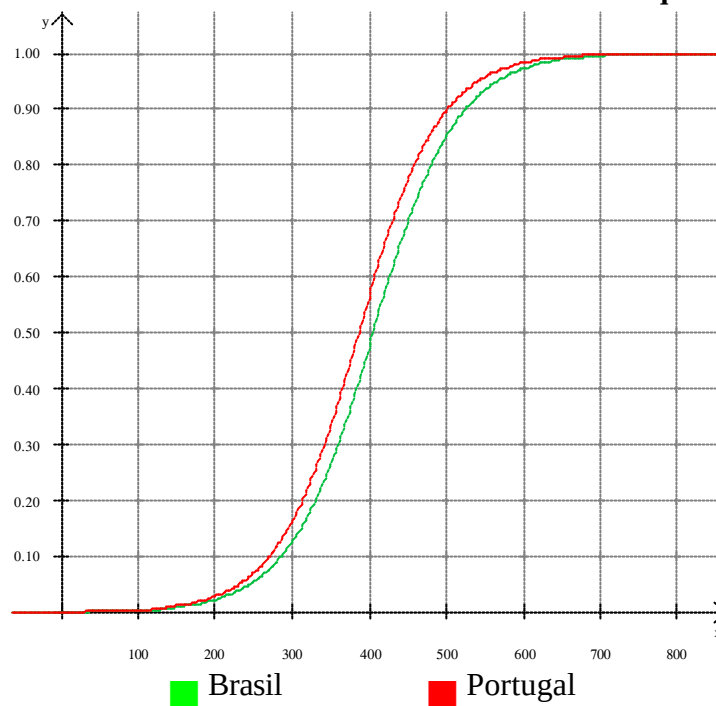
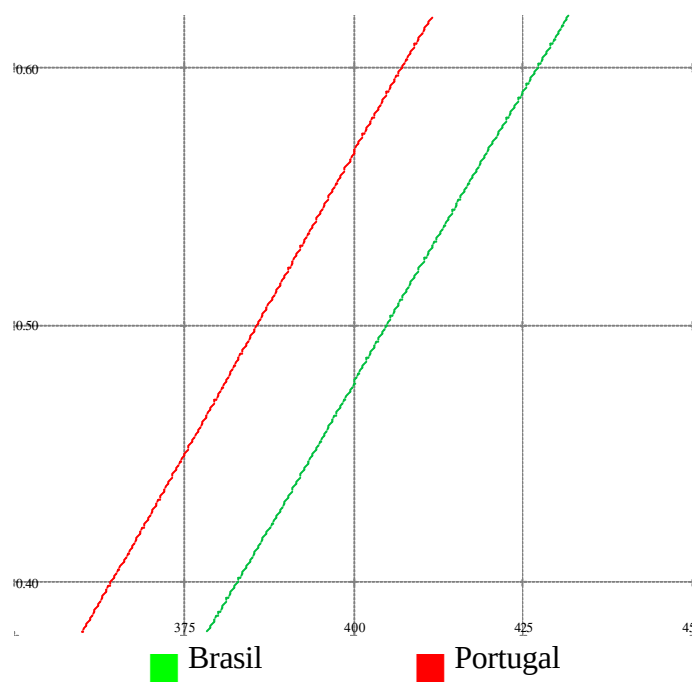
Gráfico 16: Curva Característica do Item M413q01**Gráfico 16a: Zoom da Curva Característica do Item M413q01**

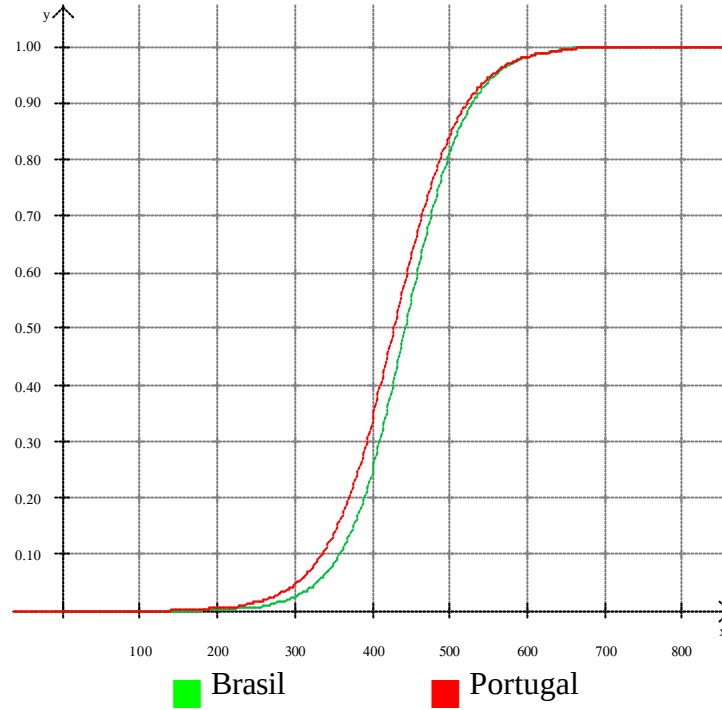
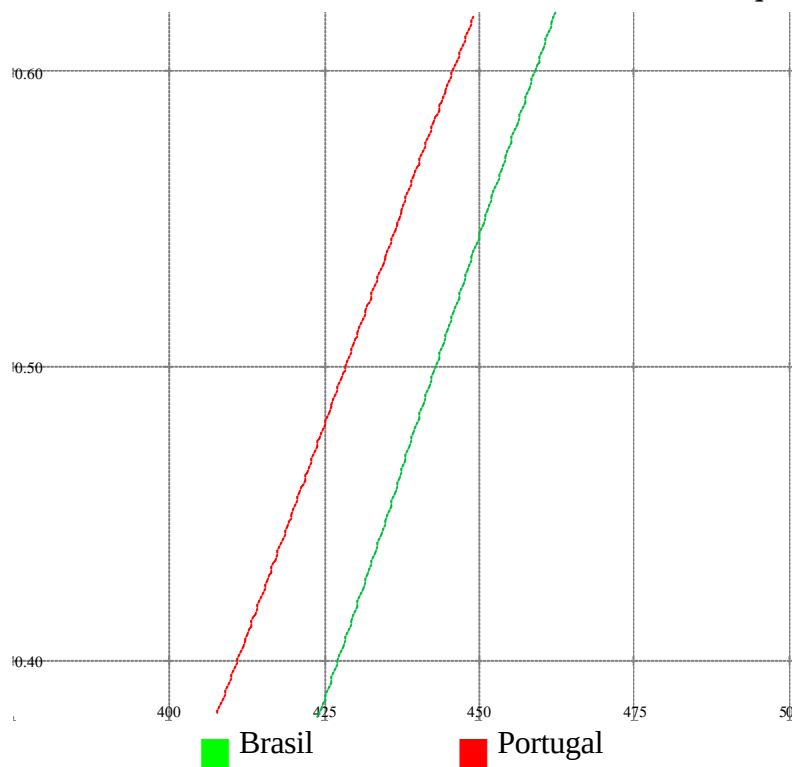
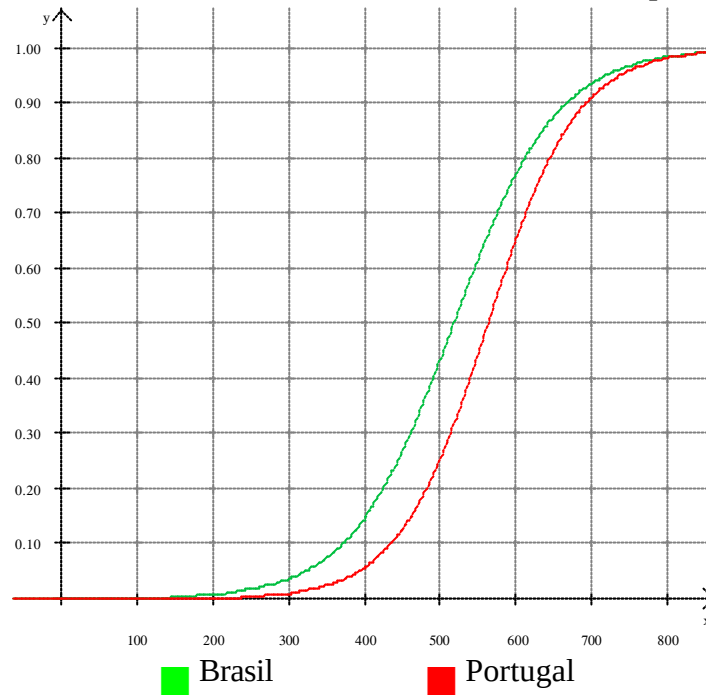
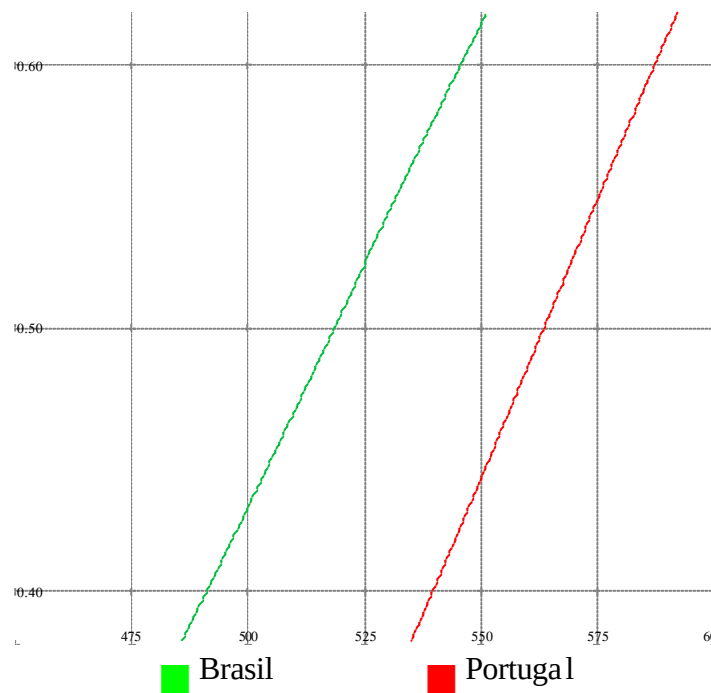
Gráfico 17: Curva Característica do Item M413q02**Gráfico 17a: Zoom da Curva Característica do Item M413q02**

Gráfico 18: Curva Característica do Item M413q03t**Gráfico 18a: Zoom da Curva Característica do Item M413q03t**

De acordo com os gráficos do ajuste do modelo pra esses três itens, apresentados a seguir, os dois primeiros mostram um melhor ajuste aos dados empíricos que o item M413Q03T. Além disso, O gráfico relativo ao percentual das escolhas (gráfico da direita) para este item mostra que ele tem o parâmetro b

maior que o dos dois itens anteriores, ou seja, é, naturalmente, um item mais difícil.

Gráfico 19: Ajuste do Modelo e CCI -Item M413Q01

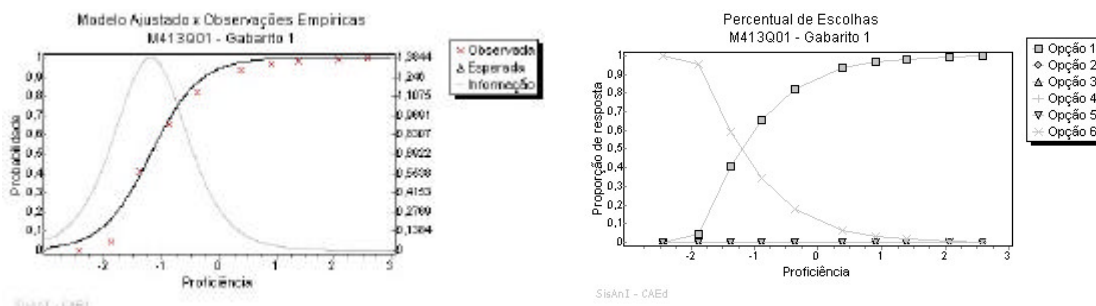


Gráfico 20: Ajuste do Modelo e CCI -Item M413Q02

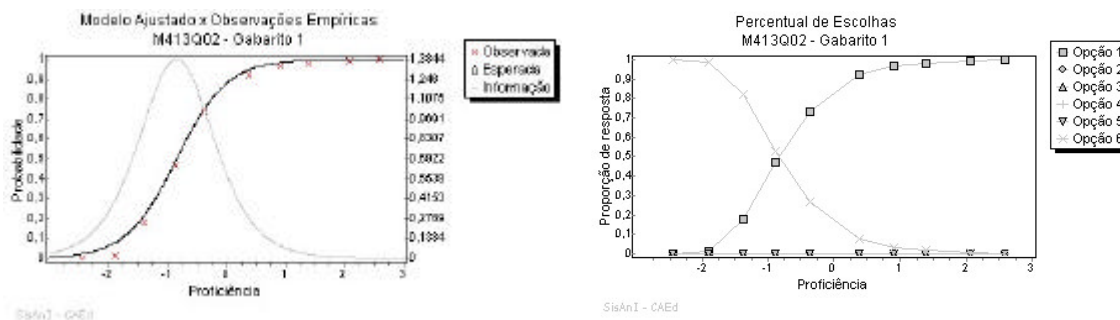
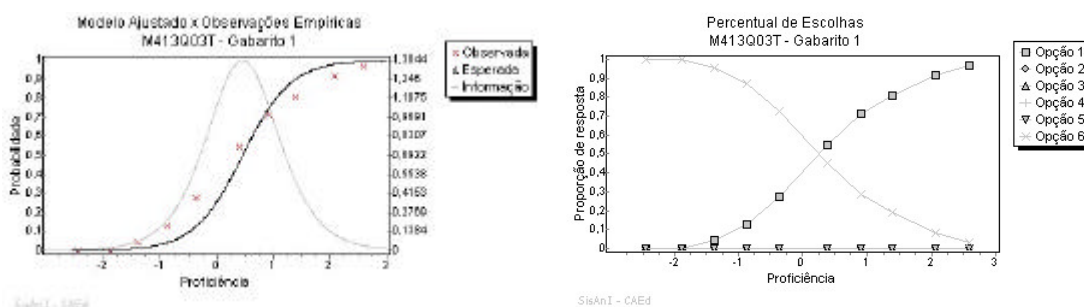


Gráfico 21: Ajuste do Modelo e CCI -Item M413Q03T

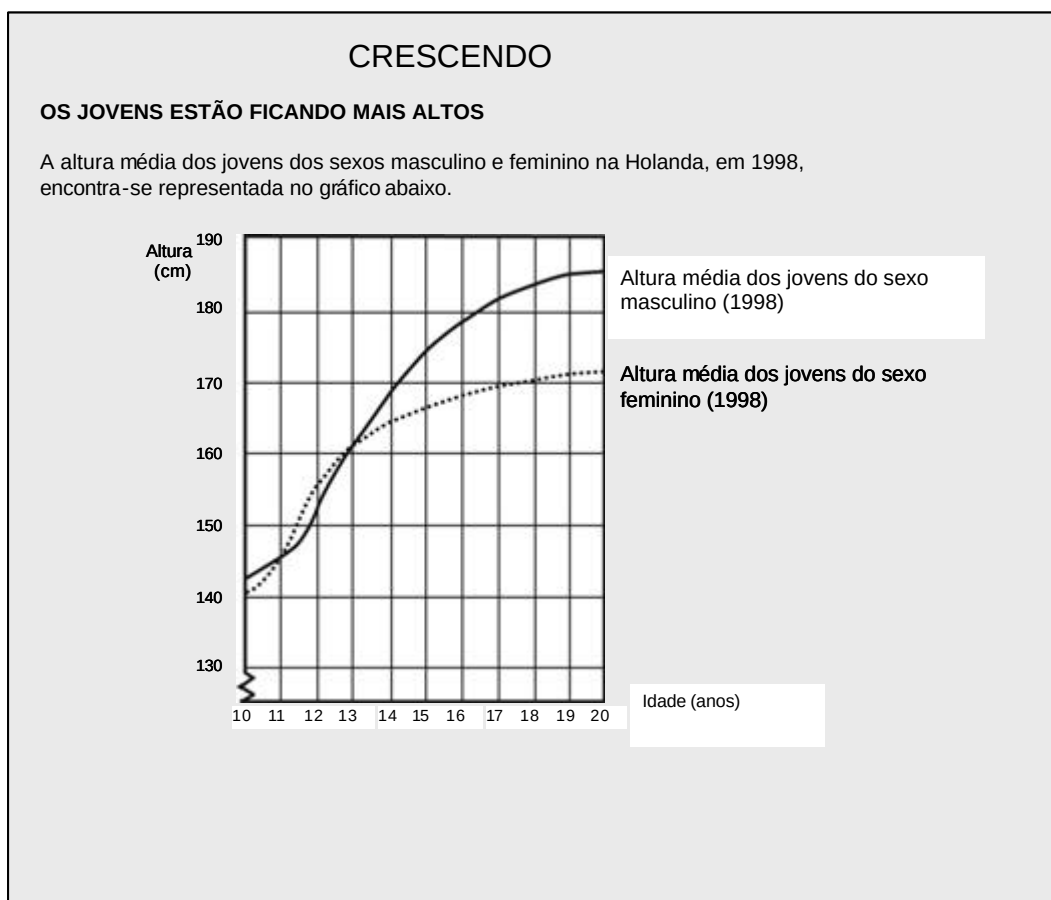


Com relação aos cinco itens da subárea “Mudança e Relações”, quatro apresentando DIF pró Portugal e um apresentando DIF pró Brasil, percebemos que em todos eles as situações problemas envolvendo manifestações matemáticas de relações e dependências funcionais entre variáveis, são feitas através da

representação gráfica. Portanto, a correta interpretação, por parte dos estudantes, dessa representação gráfica ajuda na resolução da tarefa solicitada.

Esses itens são os seguintes: M150Q03; M155Q01; M155Q02 e M302Q01T, com DIF favorável aos alunos portugueses e o item M302Q03 mais fácil para alunos brasileiros. Os itens relativos à situação problema M302 são sigilosos e não serão mostrados.

O item abaixo requer, para uma resposta correta, que o aluno domine as competências relacionadas à análise e interpretação de gráficos. Neste caso é necessário que o aluno entenda o significado da inclinação da curva que relaciona as variáveis altura e idade.



Questão 3: CRESCENDO

M150Q03 - 01 02 11 12 13 99

Explique como o gráfico permite concluir que, em média, a taxa de crescimento das meninas é mais lenta depois dos 12 anos de idade.

.....

.....

.....

A CCI para os dois grupos revela que esse item apresenta DIF também na discriminação. As curvas têm inclinações diferentes, mostrando que esse item discrimina mais os alunos brasileiros. Esse é um caso de DIF não-uniforme, havendo uma interação entre o nível de proficiência e a performance no item, ou seja, a direção do DIF muda ao longo da escala. Para alunos com proficiência abaixo de 560, aproximadamente, o item favorece os estudantes portugueses. Para alunos com proficiências maiores, o item passa a ser mais fácil para alunos brasileiros. O valor da estatística delta de Mantel Haenszel calculado para esse item ($\alpha D MH = 1,102$), indica um funcionamento diferenciado de magnitude intermediária.

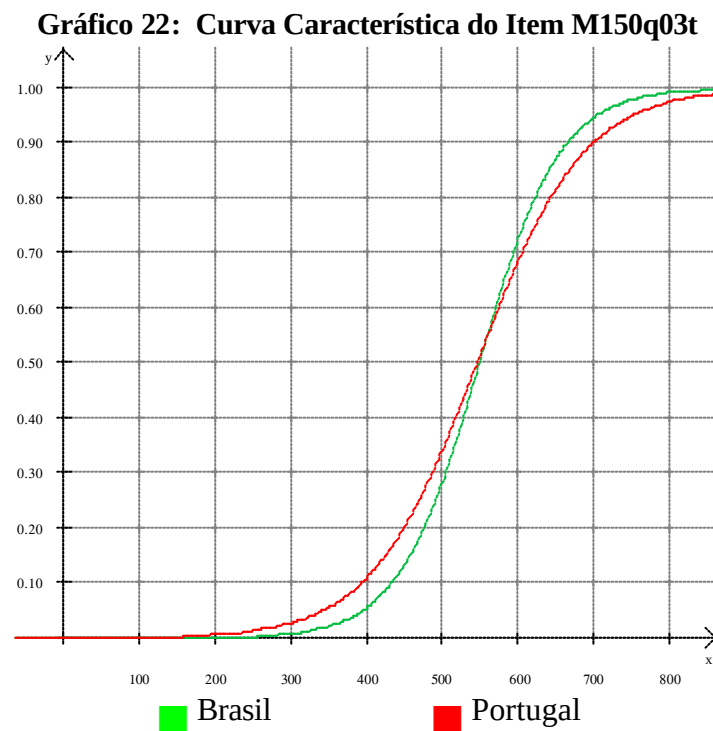
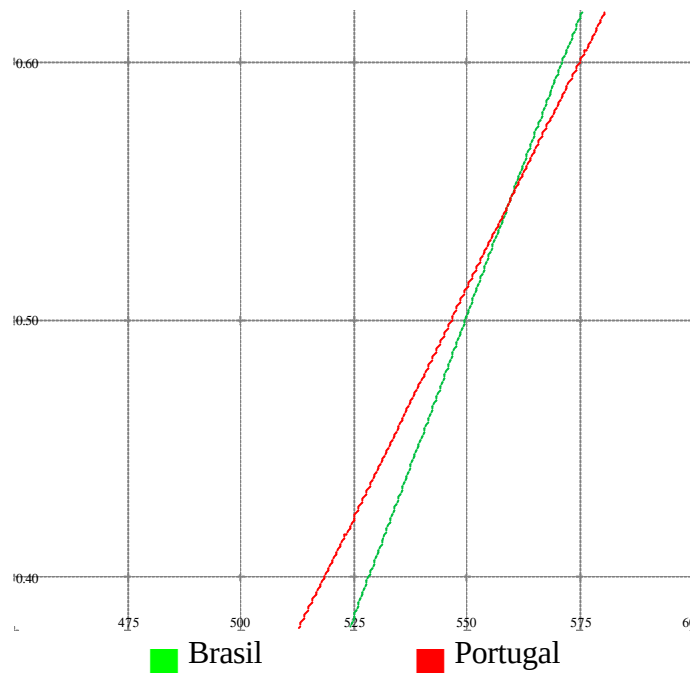
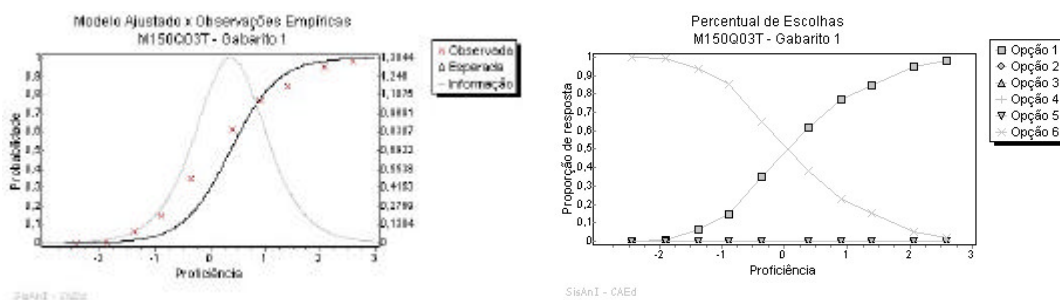


Gráfico 22a: Zoom da Curva Característica do Item M150q03t

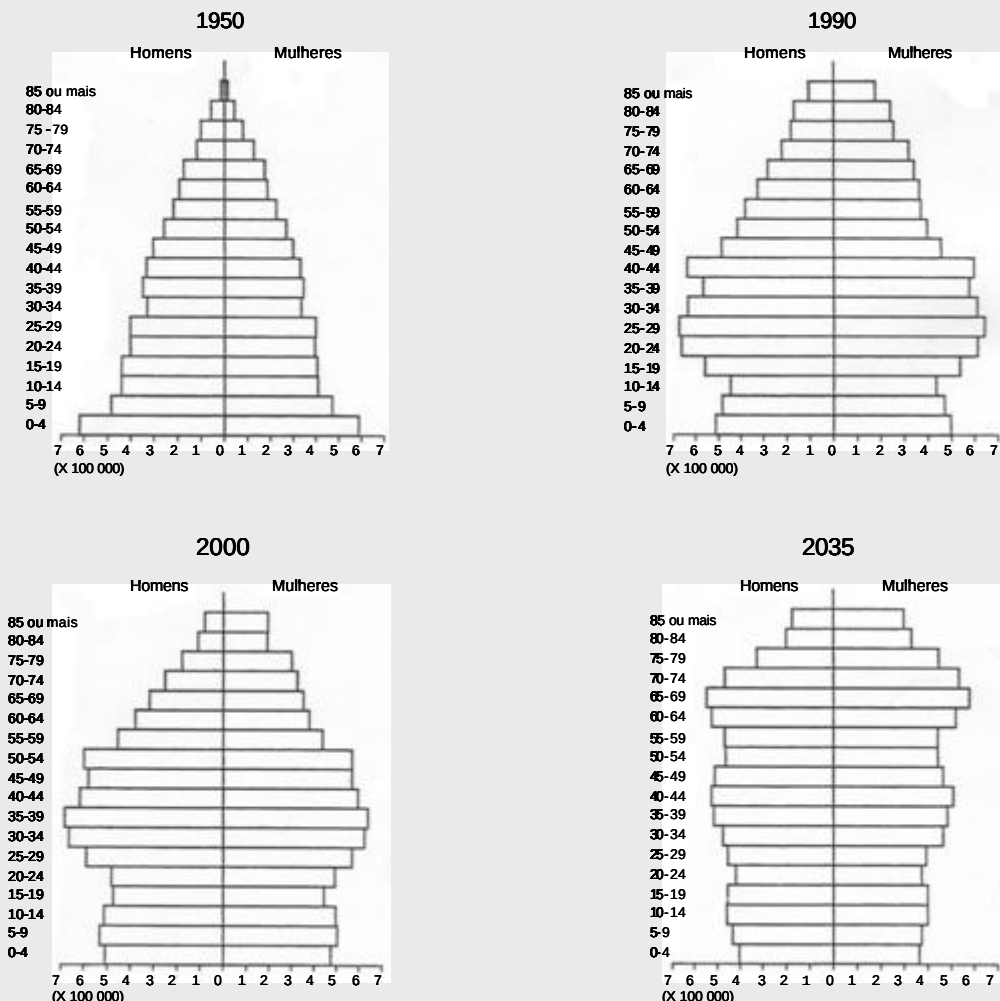
O gráfico do ajuste do modelo apresenta uma pequena subestimação dos dados empíricos para as proficiências mais baixas e superestima, ligeiramente, onde as proficiências são maiores.

Gráfico 23: Ajuste do Modelo e CCI -Item M150Q03T

Para responder corretamente os próximos itens, M155Q01 e M155Q02, o aluno necessita saber localizar pontos num gráfico e perceber a relação entre as duas variáveis (idade e gênero), em quatro pirâmides populacionais distintas. Para o segundo, ainda é exigido um cálculo estimativo.

PIRÂMIDES POPULACIONAIS

Em um certo país, para representar a distribuição populacional nos anos de 1950, 1990, 2000 e 2035, foram construídas as seguintes pirâmides populacionais:



Questão 1: PIRÂMIDES POPULACIONAIS

M155Q01-019

Em 1950, um menino chamado João tinha 3 anos de idade. Pinte com lápis as partes da pirâmide em que ele se encontra em 1950 e em 2035, supondo que ele ainda esteja vivo no ano de 2035.

Questão 2: PIRÂMIDES POPULACIONAIS

M155Q02-00 11 12 13 21 99

De acordo com as previsões, quantas pessoas (meninos e meninas juntos) na faixa etária entre 5 e 9 anos, existirão nesse país no ano de 2035? Explique como você encontrou a resposta.

.....
.....
.....

As CCIs dos itens M155Q01 e M155Q02, apresentadas abaixo, mostram que o DIF é não-uniforme, havendo interação entre a proficiência e a performance no item. Esse fato é mais acentuado no item M155Q02. Os valores da estatística delta de Mantel Haenszel calculado para esses itens são, respectivamente: $\alpha D_{MH} = 1,231$ e $\alpha D_{MH} = 1,736$. O primeiro é classificado como apresentando DIF de magnitude intermediário e o segundo de magnitude alta.

Gráfico 24: Curva Característica do Item M155q01

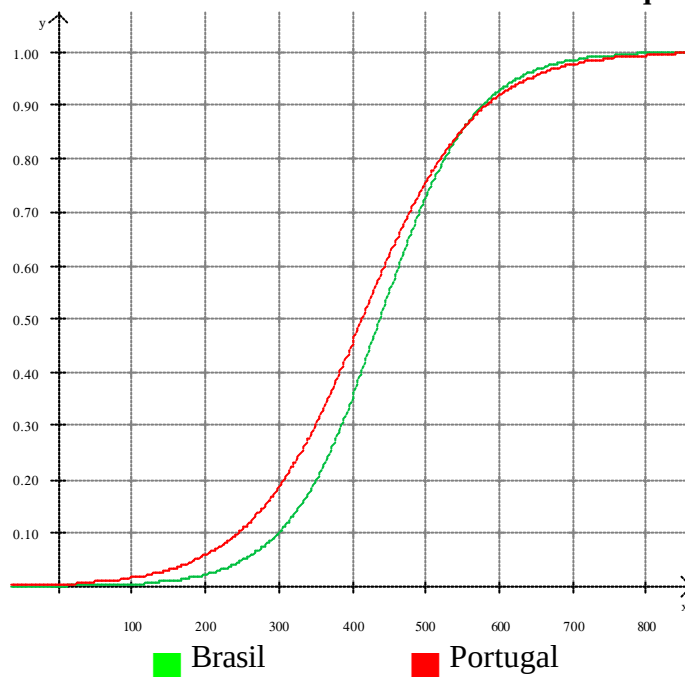


Gráfico 24a: Zoom da Curva Característica do Item M155q01

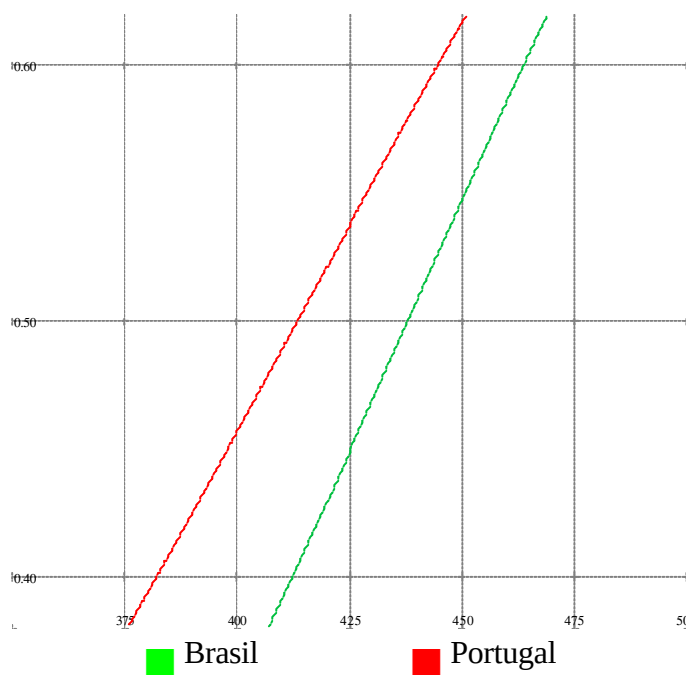
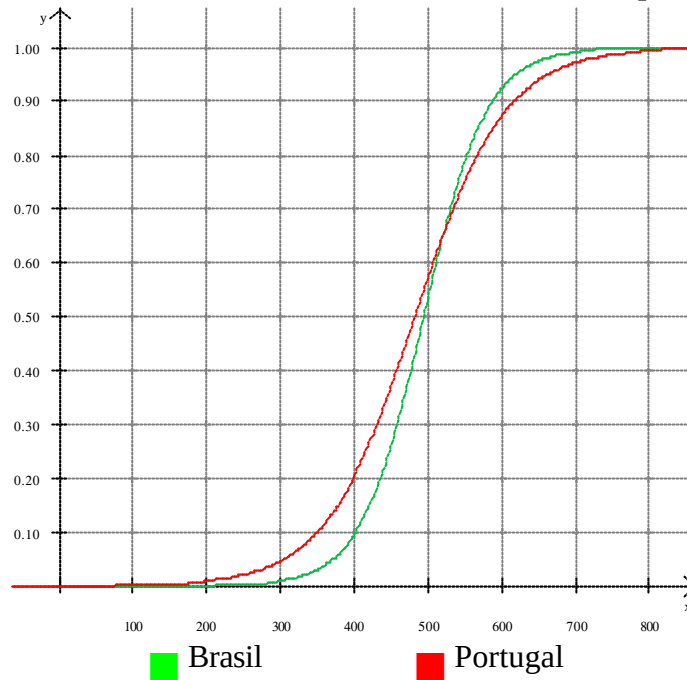
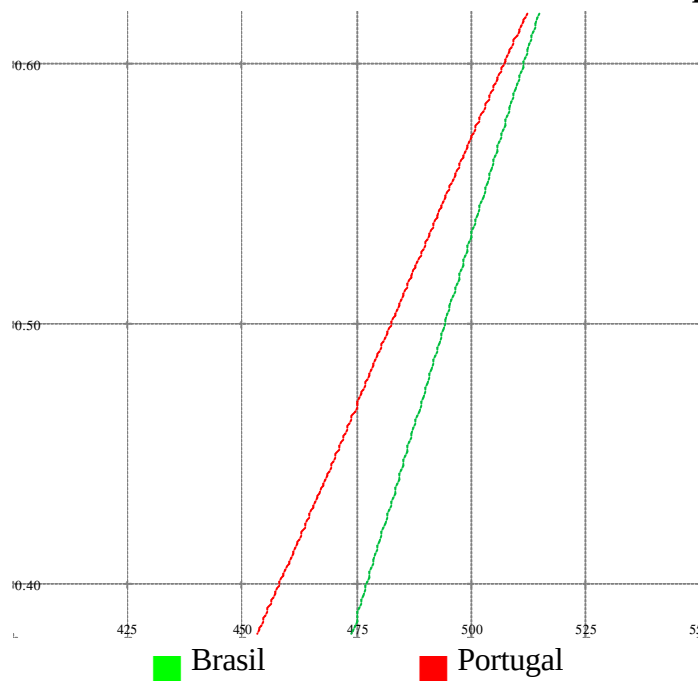
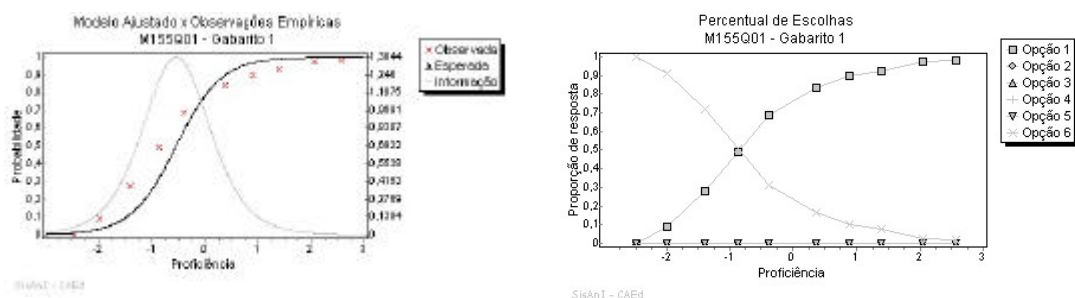
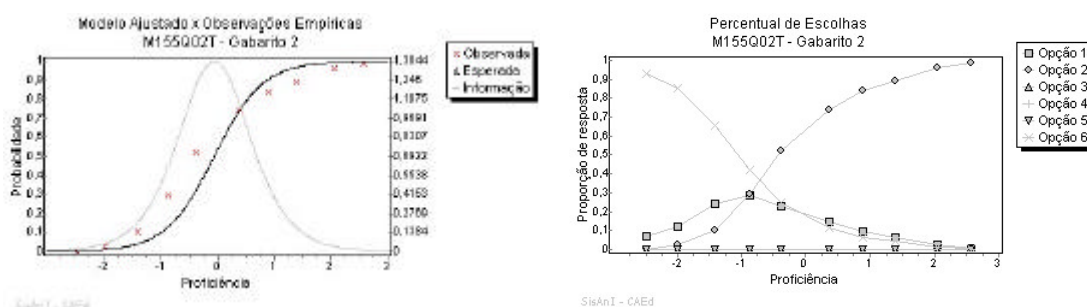


Gráfico 25: Curva Característica do Item M155q02t**Gráfico 25a: Close da Curva Característica do Item 155q02t**

De acordo com os gráficos do ajuste do modelo pra esses dois itens, apresentados a seguir, os dois mostram um ajuste razoável aos dados empíricos. Os gráficos relativos ao percentual das escolhas (gráficos da direita) para estes itens mostram um comportamento adequado em relação à opção pela resposta correta.

Gráfico 26: Ajuste do Modelo e CCI -Item M155Q01**Gráfico 27: Ajuste do Modelo e CCI -Item M155Q02T**

Nos dois itens não públicos, também de “Mudança e Relações”(um mais fácil para o Brasil e outro para Portugal), da mesma forma que os outros, a situação problema envolve relações e dependências funcionais entre variáveis através da representação gráfica, exigindo, portanto, competências e habilidades na interpretação de gráficos.

A explicação para os DIFs causados por esses itens, pode estar na maior facilidade dos alunos portugueses em estabelecer uma relação entre duas variáveis, em virtude de maior ênfase curricular na manipulação de expressões algébricas, uso de fórmulas e no uso do raciocínio proporcional. Ou ainda, no fato de que todos esses itens exigem habilidades e competência numa área que ainda pode não estar sendo adequadamente tratada no currículo ensinado em grande parte das escolas brasileiras. Estudo relacionando o perfil socioeconômico dos alunos e a seleção de conteúdos e abordagens didáticas, no Ensino Fundamental, aponta que “Tratamento da Informação”, por ser uma demanda mais recente do mundo moderno, ainda não está incorporada às práticas didáticas do ensino da Matemática (Ortigão, 2005). Embora o PISA considere a “Incerteza” como uma das quatro subáreas avaliadas, abrangendo as relações probabilísticas e fenômenos

estatísticos, muitos dos itens envolvendo gráficos são encontrados, também, na subárea “Mudança e Relações”.

5.2.2- Contexto de Aplicação: as evidências sobre o DIF associadas ao contexto “Científico” versus “Pessoal”

Na tabela abaixo, classificamos os mesmos 23 itens que apresentaram DIF em função das diferentes situações em que os estudantes encontram problemas matemáticos, ou ainda, nas situações em que são aplicados os conhecimentos relevantes. Quatro são de contexto “Científico” e os quatro tendem a apresentar DIF beneficiando o grupo focal (Portugal), indicando, portanto, serem mais fáceis para alunos portugueses. Por outro lado, quando o contexto do item é “Pessoal”, os itens que apresentam DIF, tendem a beneficiar o grupo de referência (Brasil).

Tabela 12: Itens com DIF, segundo o Contexto

Característica do Item (contexto)	Nº de itens com DIF Mais fácil para	
	Brasil	Portugal
Científico	0	4
Pessoal	4	0
Educacional/Ocupacional	3	5
Público/Social	4	3
Total	11	12

Dos quatro itens de contexto “Científico” que apresentaram DIF (são todos itens públicos) três deles são da subárea “Mudança e Relações” e já foram apresentados na seção anterior. São eles: M150q03; M155q01e M155q02t . O outro item de contexto “Científico” é M505q01 que será apresentado abaixo.

Esses itens avaliam o emprego da Matemática dentro do contexto “científico”. Em todos esses itens obteve-se o mesmo resultado: eles são mais fáceis para alunos portugueses quando comparados aos alunos brasileiros. Eles tratam de assuntos que são prioritariamente abordados no contexto escolar e, além disso, todos requerem, como já mencionado, alguma familiaridade com tipos diferentes de gráficos e suas interpretações, um tema que pertence a uma área da Matemática que parece ainda não ser adequadamente tratado em nossas salas de aula.

LIXO

Questão 1: LIXO

M505Q01 – 019

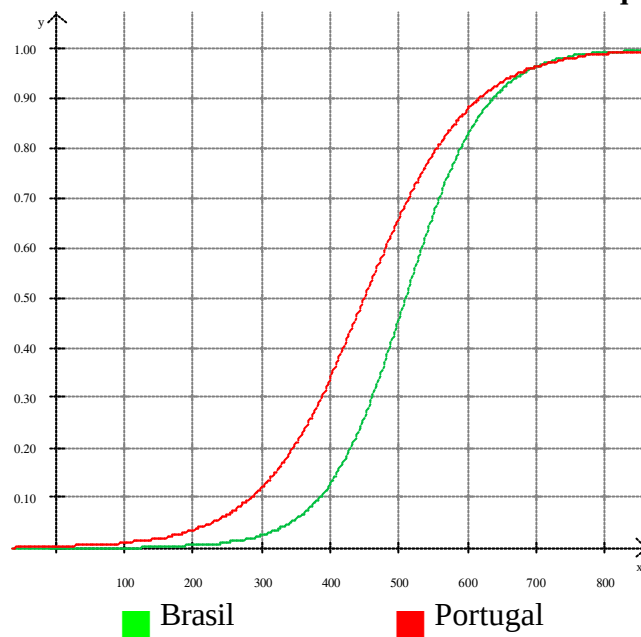
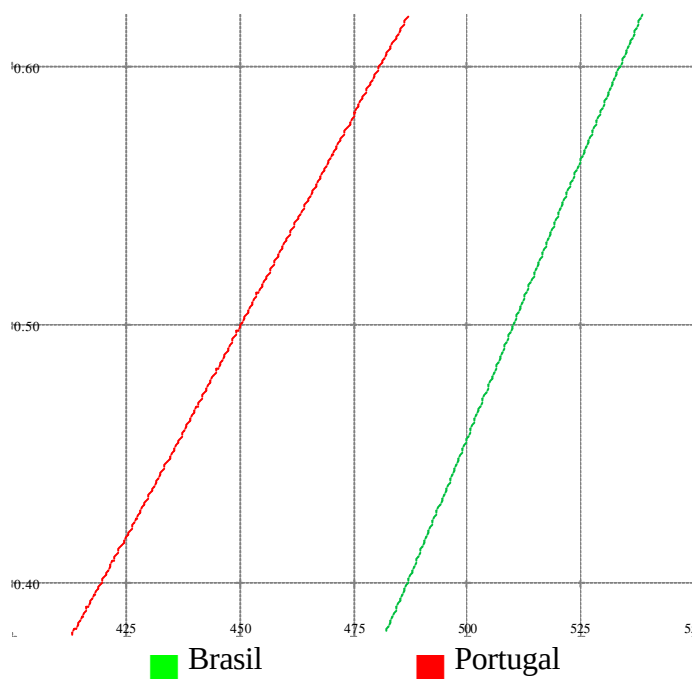
Para uma atividade escolar sobre o meio ambiente, os alunos coletaram informações sobre o tempo de decomposição de vários tipos de lixo que as pessoas jogam fora:

Tipo de lixo	Tempo de decomposição
Casca de banana	1 a 3 anos
Casca de laranja	1a 3 anos
Caixas de papelão	0,5 ano
Goma de mascar	20 a 25 anos
Jornais	Alguns dias
Copos de plástico	Mais de 100 anos

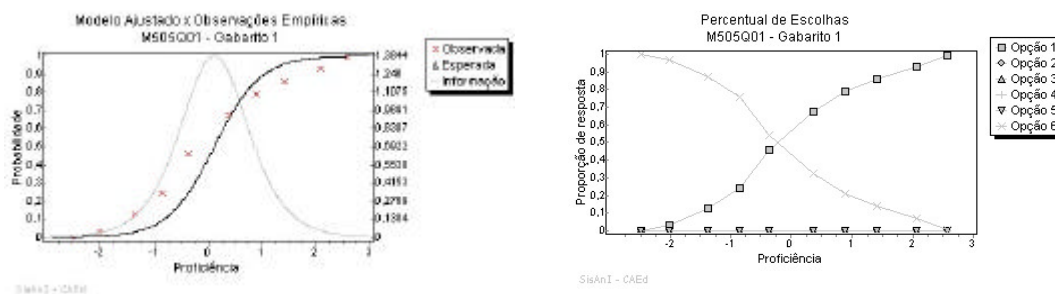
Um aluno pretende mostrar os resultados em um gráfico de barras.

Dê **uma** justificativa para o fato de que o gráfico de barras não é o mais apropriado para apresentar estes dados.

A análise da CCI deste item mostra que há DIF apenas na dificuldade, ou seja, diferem apenas no parâmetro **b** da dificuldade. Esse é um caso de DIF uniforme, pois o item favorece uniformemente os estudantes portugueses. Para todas as faixas de proficiência, o item é mais fácil para Portugal. O valor da estatística delta de Mantel Haenszel calculado para esse item (alfa D MH = 2,448), indica tratar-se de DIF de alta magnitude.

Gráfico 28: Curva Característica do Item M505q01**Gráfico 28a: Zoom da Curva Característica do Item M505q01**

Abaixo são apresentados os gráficos do ajuste do Modelo às observações empíricas e o gráfico que fornece o percentual de escolha pela opção correta em função da proficiência.

Gráfico 29: Ajuste do Modelo e CCI -Item M505Q01

Em relação ao contexto “Pessoal”, quatro itens apresentaram DIF: M496q01t; M520q02; M598q01 e M603q01t. Todos favoráveis ao Brasil, ou seja, mostraram-se mais fáceis para alunos brasileiros quando comparados com seus colegas portugueses. Desses, apenas o M520q02 é público. As situações dentro do contexto “Pessoal” são variadas estão mais próximas do cotidiano dos estudantes do que as situações de contexto “Científico”. Esses itens requerem algum tipo de competência, envolvendo contagem, que é adquirida, mais frequentemente, no contexto social e familiar do que na escola propriamente dita. Além disso, três deles são da área “Quantidade”, que, como discutido acima, tende a ser mais fácil para alunos brasileiros.

Abaixo o item M520Q02:

SKATE

Eric é um grande fã de skates. Ele vai a uma loja chamada SKATERS para verificar alguns preços.

Nesta loja, é possível comprar um skate completo, ou pode-se comprar uma prancha, um jogo de 4 rodinhas, um jogo de 2 eixos e um jogo de acessórios, e montar seu próprio skate.

Os preços dos produtos oferecidos por esta loja são:

Produto	Preço em zeds	
Skate completo	82 ou 84	
Prancha	40, 60 ou 65	
Um jogo de 4 rodinhas	14 ou 36	
Um jogo de 2 eixos	16	
Um jogo de acessórios (rolamentos, amortecedores de borracha, parafusos e porcas)	10 ou 20	

Questão 2: SKATE

M520Q02

A loja oferece três tipos de pranchas diferentes, dois jogos de rodinhas diferentes e dois jogos de acessórios diferentes. Há apenas uma opção para o jogo de eixos.

Quantos skates diferentes Eric poderá construir?

- A 6
- B 8
- C 10
- D 12

Este item já foi apresentado acima, bem como sua CCI e a estatística de Mantel Haenszel.

5.2.3- Processos: as evidências sobre o DIF associadas à “Reprodução” versus “Conexão”

A classificação dos itens, para fins de explicação do DIF, em função dos processos da Matemática que têm de ser desempenhado pelos estudantes na resolução dos itens é apresentada na tabela abaixo. Observemos as diferenças entre a quantidade de itens para cada grupo, dentro dos diferentes processos, são maiores para “Reprodução” e “Conexão”.

Tabela 13: Itens com DIF, segundo o Processo

Característica do Item (processo)	Nº de itens com DIF Mais fácil para	
	Brasil	Portugal
Reprodução	2	6
Conexão	5	2
Reflexão	4	4
Total	11	12

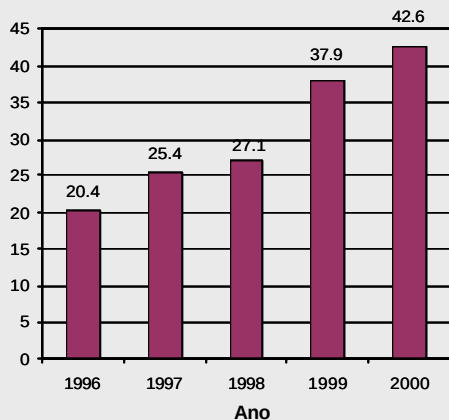
A dimensão do processo define as capacidades necessárias para a Matemática, que no caso da “Reprodução” se restringe basicamente a operações matemáticas simples. Dos oito itens que apresentaram DIF, seis foram favoráveis aos alunos portugueses. São eles: M155Q01; M155Q02; M413Q01; M413Q02; M438Q01 e M547Q01T. Apenas os itens M438Q01 e M547Q01T ainda não foram apresentados e discutidos.

Em seguida o item M438Q01:

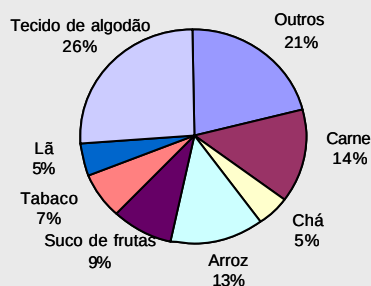
EXPORTAÇÕES

Os gráficos abaixo fornecem informações relacionadas às exportações da Zedelândia, um país que utiliza o zed como sua moeda corrente.

Total das exportações anuais da Zedelândia em milhões de zeds, 1996-2000



Distribuição das exportações da Zedelândia em 2000



Questão 1: EXPORTAÇÕES

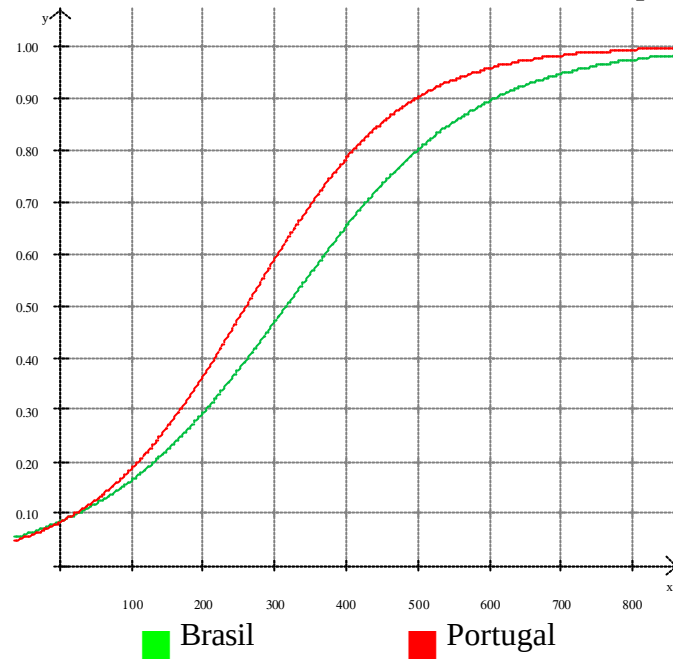
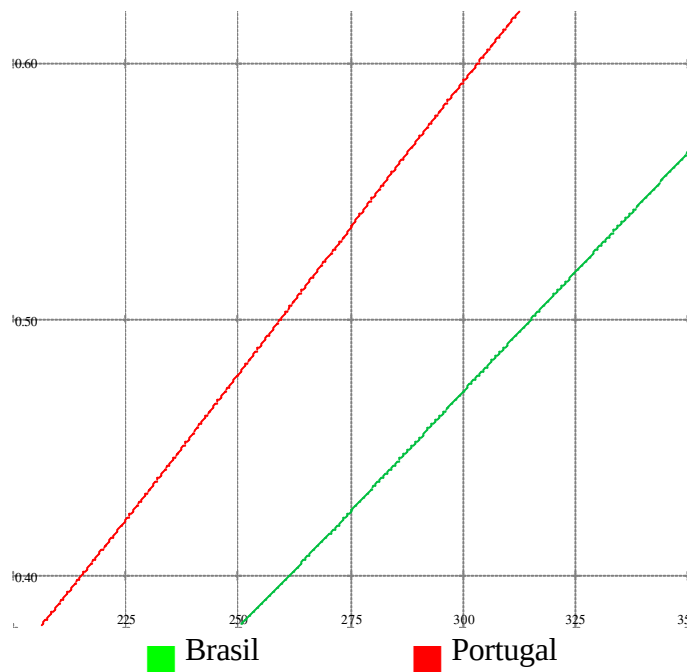
M438Q01 – 019

Qual foi o valor total (em milhões de zeds) das exportações de Zedelândia em 1998?

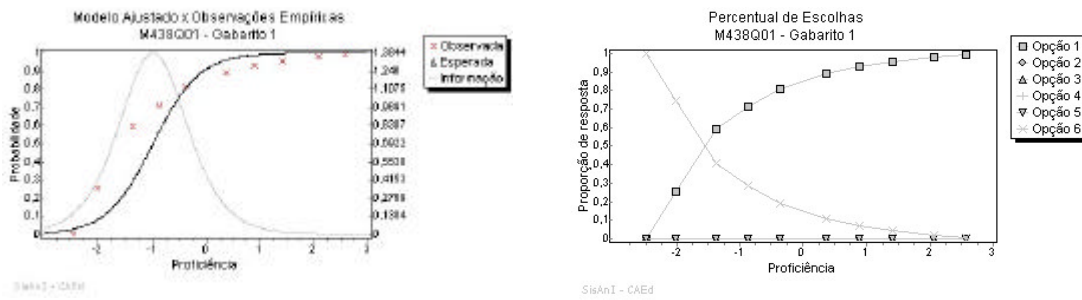
Resposta:.....

Esse item é extremamente fácil, cuja competência exigida é simplesmente localizar a informação no gráfico e reproduzi-la.

A CCI desse item, apresentadas abaixo, mostra que o DIF é uniforme, favorecendo uniformemente o grupo de alunos portugueses em relação ao grupo de alunos brasileiros. A magnitude do DIF apresentado por esse item é classificado como alta (alfa D MH = 1,514).

Gráfico 30: Curva Característica do Item M438q01**Gráfico 30a: Zoom da Curva Característica do Item M438q01**

Na sequência são apresentados os gráficos do ajuste do Modelo às observações empíricas e o gráfico que fornece o percentual de escolha pela opção correta em função da proficiência.

Gráfico 31: Ajuste do Modelo e CCI -Item M438Q01

O gráfico da esquerda mostra que o modelo subestima os dados empíricos para alunos de proficiências menores. O gráfico da direita evidencia tratar-se de um item muito fácil. Alunos com proficiências baixas já têm altas probabilidades de acertarem o item.

Outro item é o M547Q01T:

ESCADA

Questão 12: ESCADA M547Q01

A figura seguinte representa uma escada de 14 degraus, que tem uma altura total de 252 cm.

Altura total 252 cm

Profundidade total 400 cm

Qual é a altura de cada um dos 14 degraus?

Altura: cm.

Este item é da subárea “Espaço e Forma”, que requer dentre outra competências, uma percepção das formas dos objetos em representações distintas. Além disso, para responder corretamente esse item, o aluno necessita realizar uma operação de divisão com números inteiros e resto zero.

A CCI para os dois grupos revela que esse item apresenta DIF também na discriminação. As curvas têm inclinações diferentes, mostrando que esse item discrimina mais os alunos brasileiros. Esse é mais um caso de DIF não-uniforme, havendo uma interação entre o nível de proficiência e a performance no item. O valor da estatística delta de Mantel Haenszel calculado para esse item (alfa D MH = 1,506) mostra que o DIF encontrado é classificado como sendo de magnitude alta.

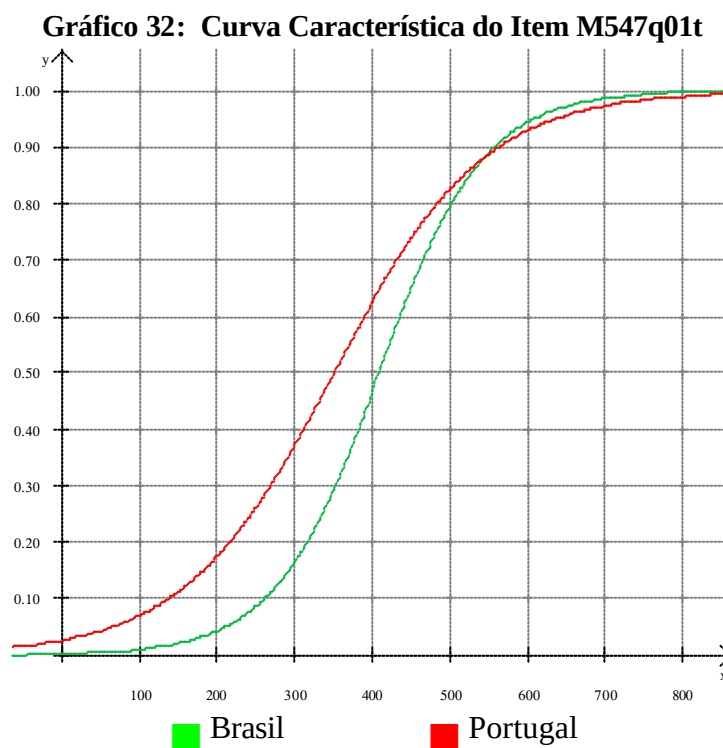
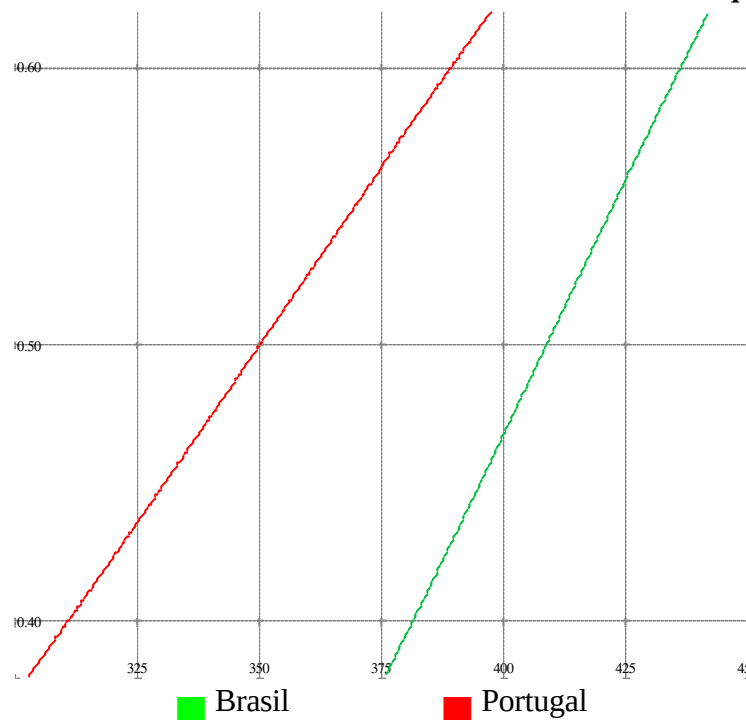
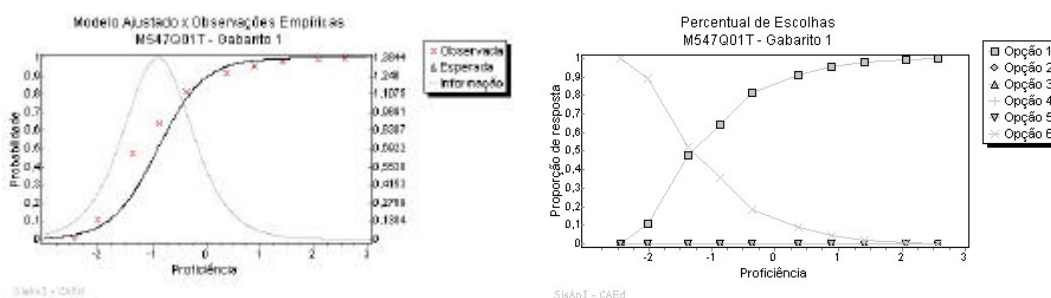


Gráfico 32a: Zoom da Curva Característica do Item M547q01t

Apresento, abaixo, os gráficos do ajuste do Modelo às observações empíricas e o gráfico que fornece o percentual de escolha pela opção correta em função da proficiência.

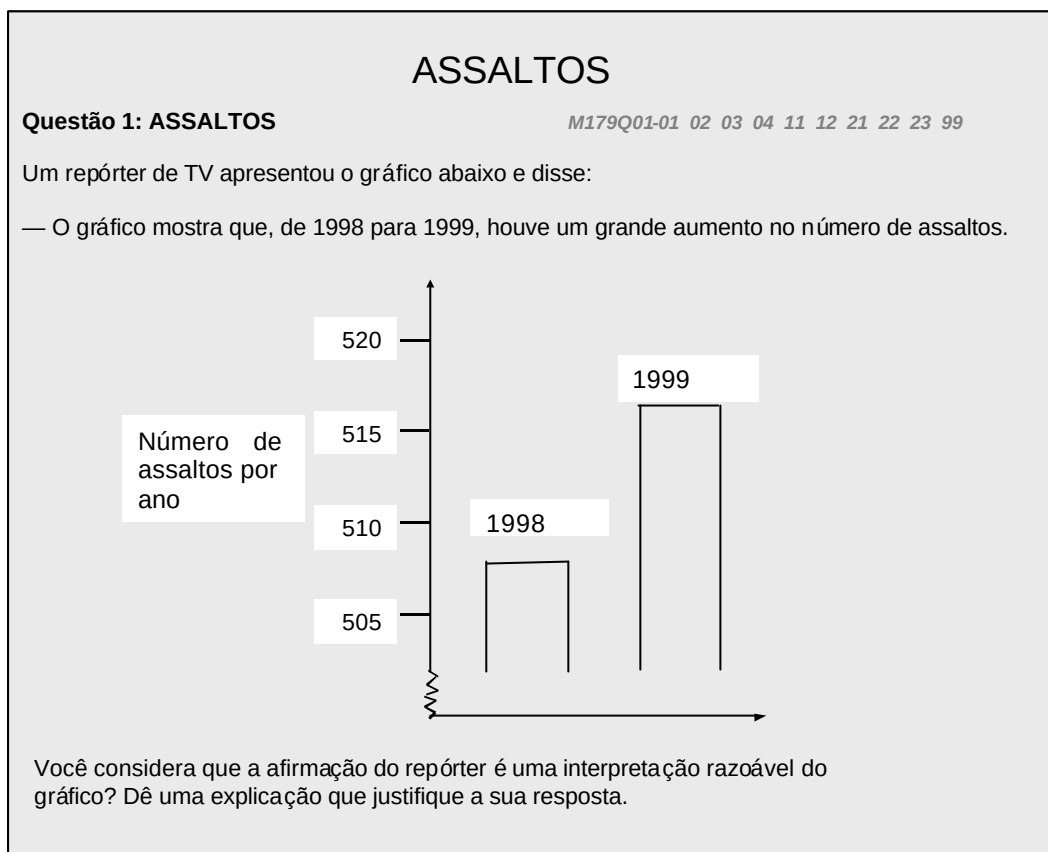
Gráfico 33: Ajuste do Modelo e CCI -Item M547Q01T

Com relação ao outro processo, a “Conexão” tem a ver com a capacidade do aluno em ligar idéias para resolver problemas de resolução direta.

Dos sete itens que apresentaram DIF, cinco foram favoráveis aos alunos brasileiros. São eles: M179Q01T; M302Q03; M438Q02; M442Q02 e M484Q01T.

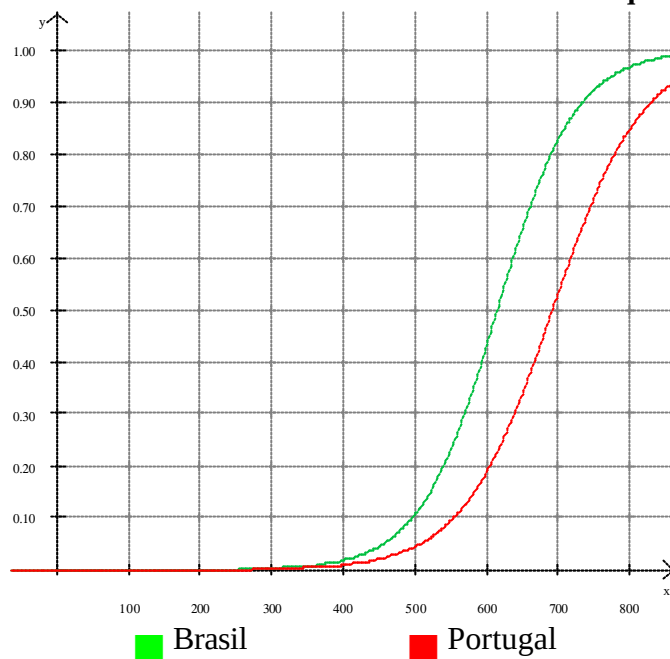
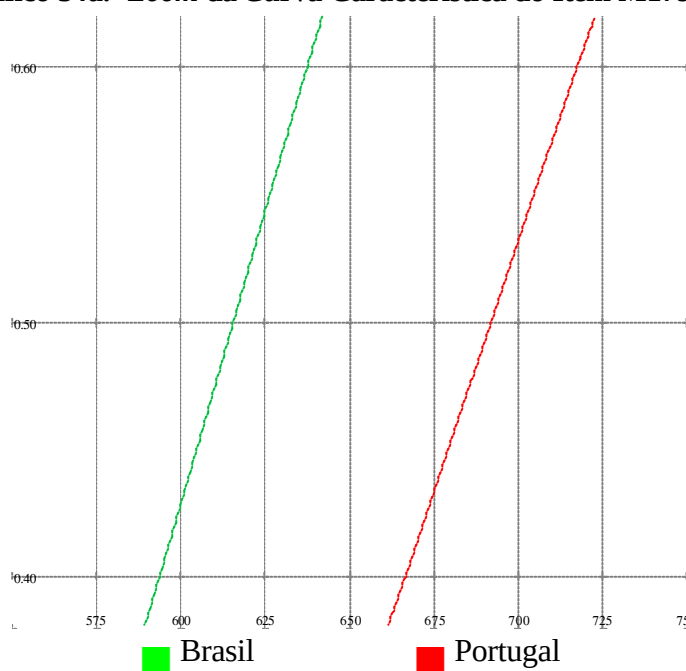
Os itens M302Q03 e M442Q02 não são itens públicos. Dos demais, o M484Q01T já foi apresentado.

Na sequência o item M179Q01T.



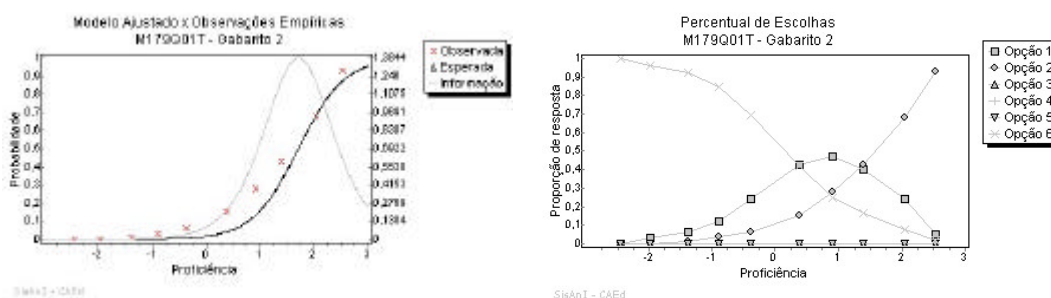
Este item, embora envolva a análise de gráfico, obrigando os alunos a refletirem sobre os valores do eixo vertical (o gráfico está incompleto), apresenta um tema que é bastante familiar aos alunos brasileiros. É razoável que para estes, um aumento no número de assaltos, de 8 ou 9, de um ano para outro, não seja considerado um “grande aumento”.

A CCI desse item, apresentadas abaixo, mostra que o DIF é uniforme, favorecendo uniformemente o grupo de alunos brasileiros em relação ao grupo de alunos portugueses. A estatística delta de Mantel Haenszel calculado para esse item (alfa D MH = -2,223), indica que o funcionamento diferenciado apresentado é classificado como de magnitude alta.

Gráfico 34: Curva Característica do Item M179q01t**Gráfico 34a: Zoom da Curva Característica do Item M179q01t**

Apresento abaixo os gráficos do ajuste do Modelo às observações empíricas e o gráfico que fornece o percentual de escolha pela opção correta em função da proficiência. O gráfico da esquerda mostra que se trata de um item relativamente difícil e que o desajuste do modelo é mais acentuado na região da curva de menores proficiências. Já o gráfico da direita mostra que somente para níveis de proficiência mais elevadas a opção correta se descola das demais opções.

Gráfico 35: Ajuste do Modelo e CCI -Item M179Q01T

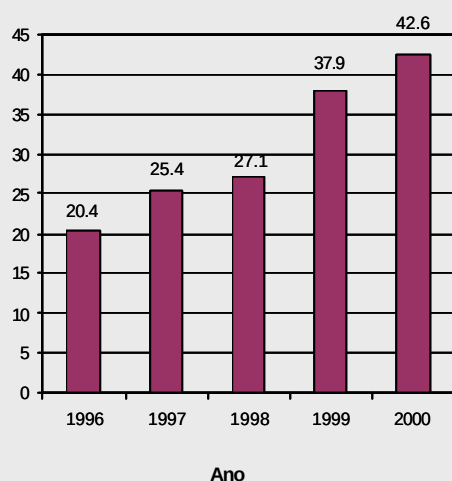


O próximo item associado ao processo “Conexão” é o item M438Q02:

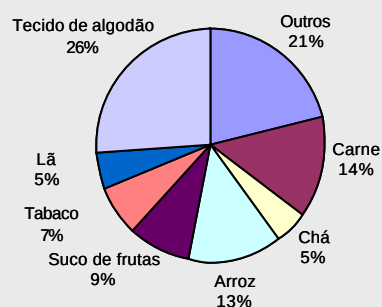
EXPORTAÇÕES

Os gráficos abaixo fornecem informações relacionadas às exportações da Zedelândia, um país que utiliza o zed como sua moeda corrente.

Total das exportações anuais da Zedelândia em milhões de zeds, 1996-2000



Distribuição das exportações da Zedelândia em 2000



Questão 2: EXPORTAÇÕES

M438Q02

Qual foi o valor total das exportações de suco de frutas de Zedelândia em 2000?

- A. 1,8 milhões de zeds.
- B. 2,3 milhões de zeds.
- C. 2,4 milhões de zeds.
- D. 3,4 milhões de zeds.
- E. 3,8 milhões de zeds.

Este item, também envolvendo análise de gráfico, exige do aluno a busca de informações nos dois gráficos apresentados e um cálculo de porcentagem. Não é um item muito fácil, embora as informações estejam bem explícitas. Como se

trata de um item de múltipla escolha, é provável que o acerto casual tenha sido o responsável pelo fato desse item ter favorecido os alunos brasileiros de baixa proficiência.

A CCI do item M438Q02 revela que esse item apresenta DIF também na discriminação. As curvas têm inclinações diferentes, mostrando uma discriminação maior para os alunos portugueses. Esse é um caso de DIF não-uniforme, havendo uma interação entre o nível de proficiência e a performance no item. O DIF encontrado é de média magnitude (alfa D MH = -1,324).

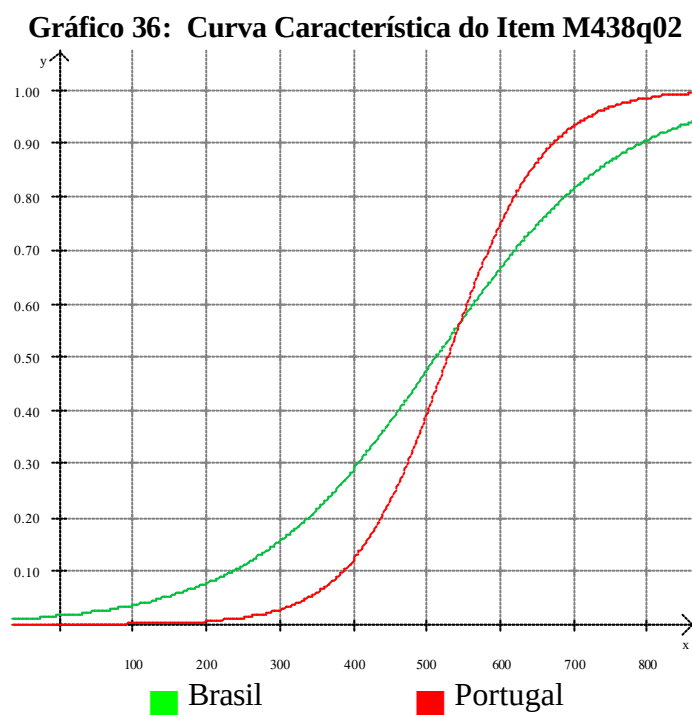
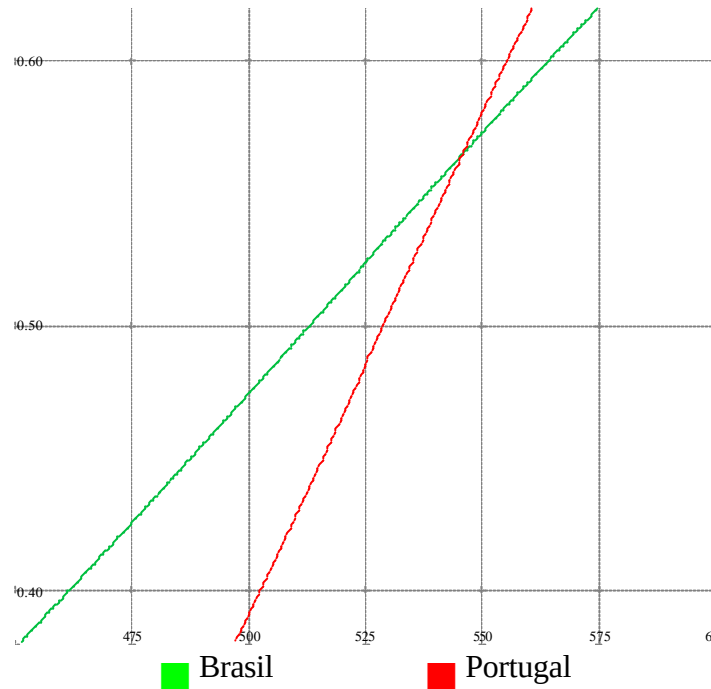
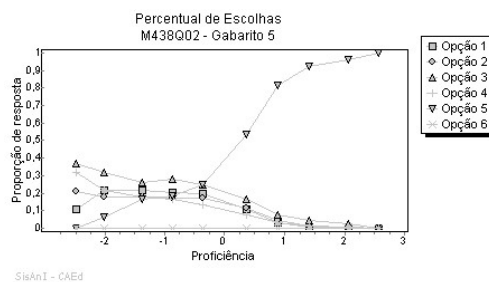
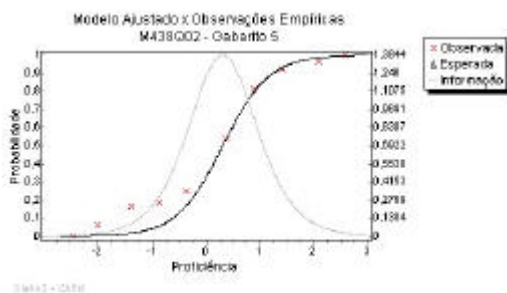


Gráfico 36a: Zoom da Curva Característica do Item M438q02

São apresentados, abaixo, os gráficos do ajuste do Modelo às observações empíricas e o gráfico que fornece o percentual de escolha pela opção correta em função da proficiência.

Gráfico 37: Ajuste do Modelo e CCI -Item M438Q02

5.2.4- Tamanho do Enunciado: as evidências sobre o DIF associadas ao “Tamanho do Enunciado”

A tabela abaixo mostra a distribuição dos 23 itens de Matemática que apresentaram algum funcionamento diferencial. A classificação dos itens, nesse momento, é feita em função do tamanho do enunciado do item.

Tabela 14: Itens com DIF, segundo o Tamanho do Enunciado

Característica do Item (enunciado)	Nº de itens com DIF Mais fácil para	
	Brasil	Portugal
Curto	4	4
Médio	3	8
Longo	4	0
Total	11	12

Observamos que, dentre os itens que apresentaram DIF, quatro deles têm a característica de possuírem um “Enunciado Longo”. São eles: M442Q02; M520Q02; M598Q01 e M603Q01. Desses, apenas o item M520Q02 é público e já foi apresentado.

Em todos esses itens o resultado foi favorável ao Brasil, ou seja, mostraram-se mais fáceis para alunos brasileiros em comparação com seus colegas portugueses de mesma habilidade cognitiva. Num primeiro momento esse resultado pode causar certa estranheza por contrariar resultados e diagnósticos anteriores, os quais nos revelam que nossos alunos lêem mal os textos que lhes são apresentados e deles conseguem extrair poucas informações para uma reflexão posterior. No entanto, ao analisar detalhadamente esses itens percebemos que o enunciado longo, na verdade, descreve instruções a serem seguidas e vêm acompanhadas de exemplos ou ilustrações indicando a tarefa a ser cumprida. Apenas o item M520Q2 não apresenta diretamente os passos a serem seguidos, acompanhado de exemplos, mas implicitamente está a idéia dos procedimentos a seguir. Provavelmente, para os alunos brasileiros, esse tipo de enunciado seja mais inteligível e as descrições dos passos a serem seguidos ajudem na concentração do que está sendo pedido e facilite a resolução do problema.

5.3- Resumo das Hipóteses Formuladas para a Explicação do DIF

Explicar DIF não é uma tarefa fácil. Além disso, toda estrutura técnica e pedagógica de uma avaliação da envergadura do PISA é empregada no sentido de se construir itens de qualidade e, principalmente, que não apresentem funcionamento diferencial entre os diferentes grupos. No entanto, sabemos que as características próprias de cada país e seu grau de desenvolvimento econômico exercem influências na vida social e cultural de sua população, com reflexos na educação. Assim, é de se esperar que os currículos sejam organizados de formas diferentes, que alguns temas sejam mais priorizados em determinados países ou que a ênfase com que se explora um determinado conteúdo seja também diferenciada. Nosso desafio para tentar explicar os DIF encontrados é verificar, nos itens que favorecem determinado país, a existência de padrões. Para isso, o ideal é que se disponha de um grande número de itens bastante diferenciados entre si, o que não foi o caso deste estudo. Um número maior de itens para esse tipo de análise talvez pudesse apontar para padrões mais claros e definidos.

No entanto, a análise descritiva das características dos itens de Matemática que apresentaram funcionamento diferencial, entre os alunos brasileiros e portugueses, mostrou-se bastante interessante. Ela revelou a existência de diferenças relacionadas à performance do item entre esses grupos e apontou caminhos para se tentar identificar padrões que possam ser causadores de DIF entre os grupos.

Sintetizo no quadro abaixo os possíveis padrões encontrados para explicar o DIF entre alunos brasileiros e portugueses. Esclareço que, da forma como foi organizado o quadro abaixo, os itens que apresentam os padrões descritos tendem a apresentar DIF favorável ao país em questão.

Quadro 36: Resumo das Características dos Itens com DIF

Grupos	Padrões
Brasil	▪ Subárea: Quantidade; ▪ Contexto: Pessoal; ▪ Processo: Conexão; ▪ Itens da subárea Quantidade, envolvendo operações algébricas básicas com números inteiros; ▪ Enunciado dos itens com orientações passo a passo e descrevendo procedimentos a serem seguidos para a resolução do item.
Portugal	▪ Subárea: Mudança e Relações; ▪ Contexto: Científico; ▪ Processo: Reprodução; ▪ Itens da Subárea Quantidade, envolvendo operações algébricas básicas com representação decimal; ▪ Itens envolvendo interpretação de variados tipos de gráficos.

Essas hipóteses formuladas para a explicação do DIF, descritas acima, carecem de uma confirmação mais rigorosa e formal que pode ser obtida a partir de uma estrutura de regressão imposta à explicação da magnitude do DIF e que tenha como co-variáveis independentes a indicação dos itens que trazem na sua forma e conteúdo as características referidas acima. Como discutido anteriormente, Soares (2007), propõe um novo modelo da TRI: uma generalização do modelo logístico de três parâmetros, que incorpora a detecção de itens com DIF em sua estrutura. Nesse modelo, além dos parâmetros a (discriminação), b (dificuldade) e c (acerto casual), são incluídos dois parâmetros - d_{ig}^a e d_{ig}^b - para representarem o comportamento diferencial na discriminação e na dificuldade, respectivamente, do item i no grupo g . O modelo proposto tem a seguinte estrutura:

$$P(Y_{ij} = 1/q_j, a_i, b_i, c_i, d_{ig}^a, d_{ig}^b) = c_i + (1 - c_i) \frac{1}{1 + e^{-De^{d_{ig}^a} a_i (q_j - b_i + d_{ig}^b)}}$$

para $i = 1, 2, \dots, I$ $j = 1, 2, \dots, J$ e $g = 1, 2, \dots, G$

Em que I representa o número total de itens, J o número total de indivíduos e G o número total de grupos.

Para um item i que não apresenta comportamento diferencial em relação à discriminação e à dificuldade, tem-se $d_{ig}^a = 0$ e $d_{ig}^b = 0$. Neste caso verifica-se que o modelo acima fica reduzido ao modelo de três parâmetros convencional:

$$P(Y_{ij} = 1/q_j, a_i, b_i, c_i) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i (q_j - b_i)}}$$

A utilização desse modelo integrado tem os seguintes objetivos:

- i) confirmar ou não as conclusões apresentadas;
- ii) verificar a concordância entre os itens detectados com DIF através das abordagens tradicionais e os detectados com DIF pela abordagem integrada.

Como dito anteriormente, espera-se obter uma maior precisão na estimação da proficiência por esta abordagem do que pela abordagem tradicional, tendo em vista que a proficiência aqui produzida não sofre o efeito dos itens que apresentam DIF.

A seguir, apresento os resultados do estudo com a utilização do modelo proposto que incorpora a detecção de itens com DIF em sua estrutura.

5.4- Resultados da Abordagem Integrada

Com o propósito de alcançar seus objetivos, acima descritos, a modelagem integrada foi implementada nos 84 itens visando, simultaneamente, detectar, estimar os parâmetros de DIF e explicar o DIF encontrado.

Considere para o estudo uma subamostra de 4496 alunos, sendo 2231 alunos brasileiros (grupo de referência) e 2265 alunos portugueses (grupo focal). As co-variáveis explicativas para o DIF que foram utilizadas no modelo são as características dos itens descritas no quadro resumo da seção 5.2, acima. Essas variáveis foram dicotomizadas para que obedecessem a seguinte estrutura: $W_{1i} = 1$ (para itens associados a uma determinada característica) e $W_{1i} = 0$ (para os demais). Assim, por exemplo, para investigar se a subárea da Matemática “Quantidade” era realmente responsável pelo DIF encontrado, foi criada uma variável, com valor igual a 1 se o item era da subárea “Quantidade” e recebeu o valor 0 caso contrário.

Primeiramente a estrutura do modelo foi implementada sem as co-variáveis explicativas. Aplicado ao conjunto de dados, o modelo proposto indicou a presença de 35 itens com DIF significativo. Um DIF definido como significativo, nesta abordagem, traduz-se por não conter o zero em seu intervalo de credibilidade ou este estar bem na extremidade do intervalo. Além disso, o critério utilizado na detecção pela abordagem integrada foi o de considerar com funcionamento diferencial uma probabilidade de apresentar DIF para um dos

grupos superior a 0.55. Considerando apenas o DIF na dificuldade, por esta abordagem, os itens com DIF significativo e positivo, ou seja, mais fáceis para Portugal (grupo focal) totalizam 25, 13 a mais do que havia sido detectado anteriormente. Os itens com DIF significativo e negativo, ou seja, mais fáceis para o Brasil (grupo de referência) totalizam 10 itens, portanto, 1 a menos em relação ao que foi detectado pela estatística de Mantel-Haenszel.

Após 15000 interações realizadas pelo modelo, obteve-se os seguintes resultados:

Tabela 15: Média e Desvio Padrão dos Grupos

Grupo	$\overline{m}_g$	$\overline{s}_g$
Brasil	-0,1003	0,81433
Portugal	0,6826	0,69497

Os resultados apresentados na tabela acima estão coerentes com os dados do PISA e as demais análises apresentadas neste trabalho, indicando que os alunos brasileiros apresentam uma proficiência, em média, bastante inferior a dos alunos portugueses. A correlação entre a proficiência do PISA e a estimada pelo modelo é de 0.830, num nível de significância 0.00.

O modelo integrado permitiu analisar o DIF com relação à discriminação e à dificuldade. Quando se analisa as estimativas encontradas para os coeficientes da estrutura de regressão dos parâmetros de DIF (g_{0g}), observa-se que o teste foi um pouco mais fácil para Portugal (valor positivo). A significância estatística, numa análise Bayesiana, pode ser constatada através do intervalo de credibilidade do parâmetro. Note-se que o intervalo de confiança associado ao coeficiente da estrutura de regressão é de [-0.0378, 0.2794], o que praticamente indica que de fato o coeficiente é significativamente positivo, já que o limite inferior do intervalo está muito próximo de zero.

Tabela 16: Coeficiente da Estrutura de Regressão do DIF (Gama) DIFICULDADE

Grupo	$\bar{g}_{0g}$	I.C (95%)	
Brasil	0	0	0
Portugal	0.1311*	-0.0378	0.2794

* Estatisticamente significativo a 95%

Quanto à discriminação, o valor de $\bar{g}_{0g}$ negativo, indica que o teste discrimina mais os alunos portugueses e o intervalo de credibilidade confirma que o coeficiente estimado é significativamente negativo.

Tabela 17: Coeficiente da Estrutura de Regressão do DIF (Gama) DISCRIMINAÇÃO

Grupo	$\bar{g}_{0g}$	I.C (95%)	
Brasil	0	0	0
Portugal	-0.1705*	-0.3332	-0.0264

* Estatisticamente significativo a 95%

O modelo proposto foi novamente aplicado, ao mesmo conjunto de dados, com a inclusão de co-variáveis indicadoras relacionadas com as características dos itens. Essas co-variáveis são as constantes do quadro 35 e foram incluídas no modelo, a princípio, individualmente. São elas:

- Quantidade e Mudança e Relações (subárea da Matemática)
- Pessoal e Científico (contexto ou situação onde o conteúdo é aplicado)
- Conexão e Reprodução (processos utilizados para resolver um problema)
- Enunciado Longo

Tabela 18: Coeficiente da Estrutura de Regressão do DIF (Gama) DIFICULDADE

Co-variáveis	$\bar{g}_{0g}$	I.C (95%)	
Quantidade	-0.2384*	-0.5598	0.0241
Pessoal	-0.1959**	-0.6257	0.1463
Conexão	-0.0766	-0.4375	0.2168
Enunciado Longo	-0.3802*	-0.7774	-0.0659
Mudança e Relações	0.1657	-0.2278	0.4946
Científico	0.2003**	-0.1538	0.5165
Reprodução	0.0334	-0.3270	0.3247

* Estatisticamente significativo a 95%

** Estatisticamente significativo a 90%

Os valores estimados para $\bar{g}_{0g}$ ratificam que itens cujas características envolvem as co-variáveis “Quantidade”, “Pessoal”, “Conexão” e “Enunciado Longo”, todas com valores negativos, são mais fáceis para alunos brasileiros. Ainda que só três delas sejam estatisticamente significativas. Já itens relacionados com “Mudança e Relações”, “Científico” e “Reprodução”, cujos respectivos valores de $\bar{g}_{0g}$ são todos positivos, são mais fáceis para alunos portugueses.

A tabela a seguir mostra as estimativas encontradas para o DIF na discriminação. Os valores negativos para “Quantidade”, “Científico” e “Reprodução”, indicam que essas co-variáveis discriminam mais os alunos portugueses. “Conexão” e “Enunciado Longo” são praticamente nulos em termos de DIF na discriminação. Já as co-variáveis “Pessoal”, e “Mudança e Relações” discriminam mais os alunos brasileiros. No entanto, são significativas apenas “Pessoal” e “Científico”.

Tabela 19: Coeficiente da Estrutura de Regressão do DIF (Gama) DISCRIMINAÇÃO

Co-variáveis	$\bar{g}_{0g}$	I.C (95%)	
Quantidade	-0.1467	-0.5035	0.1840
Pessoal	0.3648*	-0.0290	0.7575
Conexão	0.0500	-0.3082	0.4311
Enuncia do Longo	-0.0159	-0.4130	0.3520
Mudança e Relações	0.1097	-0.3816	0.6134
Científico	-0.3391*	-0.7009	0.0152
Reprodução	-0.1822	-0.6115	0.2194

* Estatisticamente significativo a 95%

Das análises acima, deve-se destacar três considerações importantes:

- i) Dos dez itens da subárea de “Quantidade” que apresentaram DIF, sete foram mais fáceis para o Brasil e três mais fáceis para Portugal (tabela 11 da seção 5.2.1). Portanto, apesar do forte indício, pertencer à subárea “Quantidade” não explica totalmente o DIF. Assim, uma análise mais detalhada mostrou que quando os itens envolviam operações algébricas básicas com números na forma decimal, a vantagem virava para o lado dos alunos portugueses;
- ii) Os quatro itens de contexto “Científico” que apresentaram DIF pró Portugal, exigiam a interpretação ou conhecimentos básicos sobre gráficos. Situação semelhante ocorreu com a subárea “Mudança e Relações”, onde as relações entre variáveis eram expressas através de gráficos. No entanto, alguns itens com DIF favoráveis ao Brasil, também exigiam algum tipo de interpretação de gráfico. No caso do contexto “Científico” parece que a explicação do DIF está mesmo no contexto, já que nenhum item com gráfico, favorável ao Brasil, era do contexto “Científico”, mas no caso de “Mudança e Relação” não se pode ter a mesma certeza;

- iii) Ao analisar os itens com “Enunciado Longo” e que apresentaram DIF favoráveis aos alunos brasileiros (contrariando algumas expectativas), percebe-se, na verdade, que o enunciado não se trata de um texto corrido. Embora com muitas palavras, a estrutura principal assemelha-se a uma instrução, passo a passo, que deve ser seguida pelo aluno para resolver o problema. Às vezes acompanhada de um exemplo, esta estrutura, ao contrário de textos prolixos, parece ajudar na concentração do aluno e discernir sobre a estratégia a adotar.

Assim, foi necessário que três novas variáveis fossem criadas e investigadas pelo modelo para a explicação do DIF: “Interpretação de Gráficos”, “Representação Decimal” e “Tipo do Enunciado”. Os resultados encontrados são apresentados nas tabelas abaixo.

**Tabela 20: Coeficiente da Estrutura de Regressão do DIF
(Gama) DIFICULDADE**

Co-variáveis	$\bar{g}_{0g}$	I.C (95%)	
Representação Decimal	0.1347	-0.1899	0.4084
Interpretação de Gráficos	-0.0637	-0.3928	0.2042
Tipo do Enunciado	-0.5530*	-1.0726	-0.1959

* Estatisticamente significativo a 95%

Note-se que, embora não significativos estatisticamente, os valores estimados de $\bar{g}_{0g}$ para essas novas co-variáveis nos permitem concluir, com relação às três considerações destacadas acima, que:

- i) Itens da subárea “Quantidade” explicam o DIF favorável aos estudantes brasileiros, em comparação com os alunos portugueses, desde que as operações aritméticas não envolvam números na representação decimal. Itens envolvendo números decimais são mais difíceis para os alunos brasileiros, comparados com seus colegas portugueses de mesma habilidade cognitiva;

- ii) É o contexto “Científico” e a subárea “Mudança e Relações” que explicam o fato de itens envolvendo essas características serem mais fáceis para alunos portugueses em comparação com os alunos brasileiros. Embora com valor estimado praticamente nulo, o sentido (valor negativo) do coeficiente $\bar{g}_{0g}$ correspondente à co-variável “Interpretação de Gráficos” aponta para o fato de que itens envolvendo a interpretação de gráficos são mais fáceis para alunos brasileiros. Essa estimativa é a única a não corroborar uma das hipóteses iniciais, ou seja, a de que itens envolvendo interpretação de variados tipos de gráficos explicam o DIF a favor de Portugal.
- iii) Selecionando dentre os itens com “Enunciado Longo” apenas aqueles onde são descritas as etapas, passo a passo, que devem ser seguidas pelos alunos para responderem corretamente o item, o resultado se confirma. Este tipo de enunciado torna os itens mais fáceis para os alunos brasileiros em comparação com os alunos portugueses.

Tabela 21: Coeficiente da Estrutura de Regressão do DIF (Gama) DISCRIMINAÇÃO

Co-variáveis	$\bar{g}_{0g}$	I.C (95%)	
Representação Decimal	0.2476*	-0.0610	0.5871
Interpretação de Gráficos	-0.1576	-0.5220	0.2002
Tipo do Enunciado	0.0422	-0.6873	0.7329

* Estatisticamente significativo a 95%

Quanto à discriminação, a co-variável “Representação decimal” é a única que aparece estatisticamente significativa e discrimina mais os alunos brasileiros do que os portugueses. Já a co-variável “Interpretação de gráficos”, embora não significativa, discrimina mais os alunos portugueses. O “Tipo do Enunciado” é praticamente nulo na explicação do DIF na discriminação.

Cabe ressaltar que, a não significância estatística em vários casos, muito se deve a pouca quantidade de itens disponíveis para este tipo de análise. À medida que selecionamos uma característica para ser investigada, o conjunto de 84 itens

fica restrito a um subconjunto muito reduzido de itens. Por exemplo, dos 84 itens da base de dados, 16 (19%) foram inicialmente classificados como tendo um “Enunciado Longo”. Mais tarde, na intenção de investigar mais detalhadamente a forma do enunciado, particularmente se eram semelhantes a instruções passo a passo, esses 16 itens ficaram reduzidos a apenas 8, que tinham essa característica. Isso muitas vezes inviabiliza a convergência das interações no modelo. Da mesma forma, não é recomendável que se analise, simultaneamente, no mesmo modelo de regressão, um número superior a três co-variáveis.

Não obstante essas limitações, um modelo mais completo, incluindo simultaneamente seis co-variáveis, foi implementado. O objetivo é investigar o efeito conjunto de algumas co-variáveis na explicação da variável dependente-DIF. As variáveis selecionadas foram aquelas que, pelas análises acima, mostraram estar associadas ao DIF favorável aos alunos brasileiros. A co-variável “Múltipla Escolha” foi introduzida em contraposição aos itens de resposta tipo “Aberta” que nas análises descritivas mostraram-se desfavoráveis aos alunos brasileiros.

As estimativas encontradas nesse modelo completo são apresentadas abaixo:

**Tabela 22: Coeficiente da Estrutura de Regressão do DIF
(Gama) DIFICULDADE**

Co-variáveis	$\bar{g}_{0g}$	I.C (95%)	
Quantidade	-0.1335**	-0.4652	0.1279
Conexão	-0.0707	-0.3857	0.1849
Pessoal	-0.0303	-0.3557	0.2426
Interpretação de Gráficos	-0.1310**	-0.4763	0.1470
Múltipla Escolha	-0.0402	-0.3843	0.2830
Tipo do Enunciado	-0.5732*	-1.2414	-0.1326
Constante	0.3502*	0.1124	0.5519

* Estatisticamente significativo a 95%

** Estatisticamente significativo a 90%

As estimativas encontradas para g_{0g} , todas negativas para o grupo de referência (Brasil), confirmam que essas co-variáveis podem, em alguma medida,

explicar o DIF favorável aos alunos brasileiros. O fato de terem sido estimadas conjuntamente faz com que haja variações em relação à significância e magnitude dos coeficientes das co-variáveis. Cabe ressaltar que o valor positivo e estatisticamente significativo da constante de regressão indica que no conjunto o teste foi mais difícil para o Brasil. Além disso, a co-variável “Interpretação de Gráficos” que estimada isoladamente não mostrou-se significativa, e com coeficiente praticamente igual a zero (-0.0637), agora aparece com coeficiente um pouco mais robusto (-0.1310) e estatisticamente significativo, para o nível de confiança de 90%. A co-variável “Quantidade”, que tinha coeficiente significativo a 95%, permanece significativo, mas agora para o nível de confiança de 90%. Destaque para a co-variável “Tipo de Enunciado”, que mostra-se consistentemente significativo, quer estimada isoladamente quer estimada neste modelo de regressão mais completo.

**Tabela 23: Coeficiente da Estrutura de Regressão do DIF
(Gama) DISCRIMINAÇÃO**

Co-variáveis	$\bar{g}_{0g}$	I.C (95%)	
Quantidade	-0.1649	-0.5731	0.2310
Conexão	0.1398	-0.2339	0.5577
Pessoal	0.1444	-0.2583	0.4974
Interpretação de Gráficos	-0.1411	-0.5561	0.2536
Múltipla Escolha	-0.0049	-0.3969	0.4118
Tipo do Enunciado	0.1222	-0.7422	0.8750
Constante	-0.0994	-0.4185	0.2329

* Estatisticamente significativo

Quanto à discriminação, os valores estimados conjuntamente para $\bar{g}_{0g}$ estão coerentes com os anteriormente apresentados para explicar o DIF na discriminação. No entanto, o valor da constante de regressão é quase nulo, indicando que, de um modo geral, as diferenças de desempenho entre alunos brasileiros e portugueses de mesma habilidade cognitiva, podem ser explicadas, na maioria dos casos, pelo DIF na dificuldade.

Neste capítulo, busquei apresentar a importância da análise do funcionamento diferencial do item na avaliação educacional. Não é difícil entender, por meio de estudos dessa natureza, que uma boa comparabilidade das proficiências estimadas para diferentes grupos de alunos avaliados depende de que haja itens comuns aplicados a esses grupos que não apresentem DIF. Por outro lado, a quase impossível tarefa de se construir instrumentos de avaliação sem DIF, para grupos muito diferentes, pode ser revertida a favor do avaliador. O DIF pode ajudar a conhecer as diferenças e esse é o primeiro passo para as ações educacionais que visam eliminá-las ou, pelo menos, diminuí-las.